GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 04:05:59 ; Search time 10554.9 Seconds
(without alignments)
1800.570 Million cell updates/sec

Title: US-09-526-329-38
Perfect score: 1152
Sequence: 1 gaggcgggtcccggttgcag.....aaaaaaaaaaaaaaaaaaaa 1152

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1152	100.0	1152	9	AF110466	AF110466 Homo sapi
2	1150.4	99.9	1177	9	AF137027	AF137027 Homo sapi
3	963	83.6	3144	9	AB035343	AB035343 Homo sapi
4	963	83.6	3240	9	AB035342	AB035342 Homo sapi
5	963	83.6	3532	9	AB035340	AB035340 Homo sapi
6	963	83.6	3599	9	AB035341	AB035341 Homo sapi
c 7	718	62.3	114438	2	AC010356	AC010356 Homo sapi
c 8	718	62.3	204843	9	AC010359	AC010359 Homo sapi
9	701.6	60.9	137808	2	AC008889	AC008889 Homo sapi
10	701.2	60.9	6486	9	AF110465	AF110465 Homo sapi
11	700.4	60.8	1248	9	AF110467	AF110467 Homo sapi
12	699.6	60.7	166308	9	CNS01DX2	AL139020 Human chr
13	699.6	60.7	210791	9	CNS01DV7	AL133467 Human chr
14	695.4	60.4	110879	9	AC016603	AC016603 Homo sapi
15	687.6	59.7	1717	9	AB025272S3	AB025274 Homo sapi
16	687.2	59.7	1722	9	AB018563	AB018563 Homo sapi
17	189	16.4	627	9	AB025272S1	AB025272 Homo sapi
18	169.4	14.7	753	9	AB025272S2	AB025273 Homo sapi
19	64.4	5.6	7218	6	I66494	I66494 Sequence 14
20	63	5.5	1040	10	AF195492	AF195492 Mus muscu
21	62.8	5.5	1231	9	BC003574	BC003574 Homo sapi
22	62.8	5.5	1240	9	BC005831	BC005831 Homo sapi
23	62.8	5.5	1324	6	AR086519	AR086519 Sequence
24	62.8	5.5	1324	9	HSTCL1	X82240 H.sapiens m
25	56.6	4.9	1048	10	AF195493	AF195493 Mus muscu
26	55	4.8	1020	10	AF195489	AF195489 Mus muscu
27	51.2	4.4	1058	10	AF195490	AF195490 Mus muscu
c 28	51	4.4	166308	9	CNS01DX2	AL139020 Human chr
29	50	4.3	4922	6	AR086521	AR086521 Sequence
30	48	4.2	1338	10	AF031956	AF031956 Mus muscu
31	48	4.2	1338	10	MMRNATCL1	Y15376 Mus musculu
32	47.4	4.1	957	10	AF195491	AF195491 Mus muscu
c 33	45.6	4.0	10732	6	E32986	E32986 Gene encodi
c 34	44.2	3.8	124499	2	AC007307	AC007307 Mus muscu
c 35	42.6	3.7	134148	2	AL358512	AL358512 Homo sapi
36	42	3.6	1040	10	AF195488	AF195488 Mus muscu
37	41.6	3.6	1719	6	AX127641	AX127641 Sequence
38	41.2	3.6	1770	5	CCZP2A	Z72491 C.carpio mR
39	41	3.6	1555	9	BC001141	BC001141 Homo sapi
40	41	3.6	1555	9	BC008033	BC008033 Homo sapi
41	40.8	3.5	89723	9	AP001417	AP001417 Homo sapi
42	40.8	3.5	100000	9	AP000018	AP000018 Homo sapi
43	40.8	3.5	100000	9	AP000160	AP000160 Homo sapi
44	40.8	3.5	340000	9	AP001730	AP001730 Homo sapi
c 45	40.4	3.5	179463	2	AL596453	AL596453 Homo sapi

ALIGNMENTS

RESULT 1
AF110466
LOCUS AF110466 1152 bp mRNA PRI 18-MAR-1999
DEFINITION Homo sapiens T-cell leukemia/lymphoma 1B (TCL1B) mRNA, complete cds.
ACCESSION AF110466
VERSION AF110466.1 GI:4324704
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1152)
AUTHORS Pekarsky,Y., Hallas,C., Isobe,M., Russo,G. and Croce,C.M.
TITLE Abnormalities at 14q32.1 in T cell malignancies involve two oncogenes
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2949-2951 (1999)
MEDLINE 99178995
REFERENCE 2 (bases 1 to 1152)
AUTHORS Pekarsky,Y., Hallas,C. and Croce,C.M.

```
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-1998) KCI, Thomas Jefferson University, 233 South
            10th Str. BLSB 1032, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .1152
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /map="14q32.1"
     gene            1. .1152
                     /gene="TCL1B"
     CDS             28. .414
                     /gene="TCL1B"
                     /codon_start=1
                     /product="T-cell leukemia/lymphoma 1B"
                     /protein_id="AAD16997.1"
                     /db_xref="GI:4324705"
                     /translation="MASEASVRLGVPPGRLWIQRPGIYEDEEGRTWVTVVVRFNPSRR
                     EWARASQGSRYEPSITVHLWQMAVHTRELLSSGQMPFSQLPAVWQLYPGRKYRAADSS
                     FWEIADHGQIDSMEQLVLTYQPERKD"
BASE COUNT      232 a    325 c    324 g    271 t
ORIGIN


  Query Match             100.0%;  Score 1152;  DB 9;  Length 1152;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1152;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gaggcgggtcccggttgcagacttgccatggcctccgaagcttctgtgcgtctaggggtg 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GAGGCGGGTCCCGGTTGCAGACTTGCCATGGCCTCCGAAGCTTCTGTGCGTCTAGGGGTG 60

Qy      61 ccccctggccgtctgtggatccagaggcctggcatctacgaagatgaggaggggagaacc 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CCCCCTGGCCGTCTGTGGATCCAGAGGCCTGGCATCTACGAAGATGAGGAGGGGAGAACC 120

Qy     121 tgggtgactgtggtcgtgcggttcaatccctcgcgtagggaatgggccagggcctcccag 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGGGTGACTGTGGTCGTGCGGTTCAATCCCTCGCGTAGGGAATGGGCCAGGGCCTCCCAG 180

Qy     181 ggcagcagatatgaacccagcatcacagtgcacttgtggcagatggcagtgcatacccgg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GGCAGCAGATATGAACCCAGCATCACAGTGCACTTGTGGCAGATGGCAGTGCATACCCGG 240

Qy     241 gagctactctcctccggccagatgcccttctcccagctgcccgccgtgtggcagctctac 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GAGCTACTCTCCTCCGGCCAGATGCCCTTCTCCCAGCTGCCCGCCGTGTGGCAGCTCTAC 300

Qy     301 cccgggaggaagtaccgagcagcggattccagtttctgggaaatagcagaccatggccag 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CCCGGGAGGAAGTACCGAGCAGCGGATTCCAGTTTCTGGGAAATAGCAGACCATGGCCAG 360

Qy     361 attgactctatggagcagctggtcctaacatatcagccggagaggaaagactgacactgg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ATTGACTCTATGGAGCAGCTGGTCCTAACATATCAGCCGGAGAGGAAAGACTGACACTGG 420

Qy     421 gagtggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcag 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAGTGGCTGGCCCTGCTGGCCCTGCCTCTTCTGGCCTGGTGTCTCCTCATGCCCCCTCAG 480

Qy     481 tgaggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggca 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGAGGATCTTCATGTACCTGCTCTTCTGTTTGCACACCCAGCATAGCCTCCTTGCAGGCA 540

Qy     541 gaaggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtcctgt 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GAAGGCAGTAGGGCCCCTGCACACTCAGTTTCTCTCGTTTTCCTTAGTTATCAGTCCTGT 600

Qy     601 cctgtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggccct 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CCTGTCCCACTCAGGTCTGTACTTAGGGCAGCTGGCCTGGATGGGCTTCACTGGGGCCCT 660

Qy     661 gtctgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgt 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTCTGTGTGCTGAGCCAGTTTCCCCTGCTGGCTGCAAGCTGTGGGTTCTTTCTCCTCTGT 720

Qy     721 gcccctcatgctgatcttctagatgccactcccaaatccccttcatacccaccaggatgt 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GCCCCTCATGCTGATCTTCTAGATGCCACTCCCAAATCCCCTTCATACCCACCAGGATGT 780

Qy     781 gtgcccagccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggc 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GTGCCCAGCCAGGCCTCCAGCACCCCCAGTGCAGCTCGTGATTGGAAACTCACCATCGGC 840

Qy     841 aggcagtggttcggtttaagagatggcattagagggagcccagtctggatgtggacttgg 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 AGGCAGTGGTTCGGTTTAAGAGATGGCATTAGAGGGAGCCCAGTCTGGATGTGGACTTGG 900

Qy     901 atgccctgtgggtatcagttctgctgacactttggcccgaaatagatccagtgctgagca 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ATGCCCTGTGGGTATCAGTTCTGCTGACACTTTGGCCCGAAATAGATCCAGTGCTGAGCA 960

Qy     961 agcaatgtacaccggagcctcagtgagcccatctgcacagtggggagcatggagggatgg 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 AGCAATGTACACCGGAGCCTCAGTGAGCCCATCTGCACAGTGGGGAGCATGGAGGGATGG 1020

Qy    1021 gtttggcctgtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtg 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GTTTGGCCTGTGCTTCTGCTTATTCAGTCCTTCAGCTCACGGAAGGGATGCTAGTCCGTG 1080

Qy    1081 aaggtgacctcacagtactggttaattaaacttttattgctcactgtcaaaaaaaaaaaaa 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 AAGGTGACCTCACAGTACTGGTTAATTAAACTTTATTGCTCACTGTCAAAAAAAAAAAAA 1140

Qy    1141 aaaaaaaaaaaa 1152
           ||||||||||||
Db    1141 AAAAAAAAAAAA 1152


RESULT   2
AF137027
LOCUS       AF137027     1177 bp    mRNA            PRI       13-MAY-1999
DEFINITION  Homo sapiens syncytiotrophoblast-specific protein mRNA, complete
            cds.
ACCESSION   AF137027
VERSION     AF137027.1  GI:4809182
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1177)
  AUTHORS   Jiang,B. and Mendelson,C.R.
  TITLE     A syncytiotrophoblast-specific gene Syn-1 cloned from human
            syncytiotrophoblast subtracted cDNA library
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1177)
  AUTHORS   Jiang,B. and Mendelson,C.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-1999) Biochemistry, UT Southwestern Medical
            Center at Dallas, 5323 Harry Hines Blvd., Dallas, Texas 75235, USA
FEATURES             Location/Qualifiers
     source          1. .1177
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             47. .433
                     /note="Syn-1 protein"
                     /codon_start=1
                     /product="syncytiotrophoblast-specific protein"
                     /protein_id="AAD30130.1"
                     /db_xref="GI:4809183"
                     /translation="MASEASVRLGVPPGRLWIQRPGIYEDEEGRTWVTVVVRFNPSRR
                     EWARASQGSRYEPSITVHLWQMAVHTRELLSSGQMPFSQLPAVWQLYPGRKYRAADSS
                     FWEIADHGQIDSMEQLVLTYQPERKD"
BASE COUNT      242 a    328 c    332 g    275 t
ORIGIN
```

```
  Query Match            99.9%;  Score 1150.4;  DB 9;  Length 1177;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 1151;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 gaggcgggtcccggttgcagacttgccatggcctccgaagcttctgtgcgtctaggggtg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      20 GAGGCGGGTCCCGGTTGCAGACTTGCCATGGCCTCCGAAGCTTCTGTGCGTCTAGGGGTG 79

Qy      61 ccccctggccgtctgtggatccagaggcctggcatctacgaagatgaggaggggagaacc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      80 CCCCCTGGCCGTCTGTGGATCCAGAGGCCTGGCATCTACGAAGATGAGGAGGGGAGAACC 139

Qy     121 tgggtgactgtggtcgtgcggttcaatccctcgcgtagggaatgggccagggcctcccag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 TGGGTGACTGTGGTCGTGCGGTTCAATCCCTCGCGTAGGGAATGGGCCAGGGCCTCCCAG 199

Qy     181 ggcagcagatatgaacccagcatcacagtgcacttgtggcagatggcagtgcatacccgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 GGCAGCAGATATGAACCCAGCATCACAGTGCACTTGTGGCAGATGGCAGTGCATACCCGG 259

Qy     241 gagctactctcctccggccagatgcccttctcccagctgcccgccgtgtggcagctctac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 GAGCTACTCTCCTCCGGCCAGATGCCCTTCTCCCAGCTGCCCGCCGTGTGGCAGCTCTAC 319

Qy     301 cccgggaggaagtaccgagcagcggattccagtttctgggaaatagcagaccatggccag 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     320 CCCGGGAGGAAGTACCGAGCAGCGGATTCCAGTTTCTGGGAAATAGCAGACCATGGCCAG 379

Qy     361 attgactctatggagcagctggtcctaacatatcagccggagaggaaagactgacactgg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     380 ATTGACTCTATGGAGCAGCTGGTCCTAACATATCAGCCGGAGAGGAAAGACTGACACTGG 439

Qy     421 gagtggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcag 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     440 GAGTGGCTGGCCCTGCTGGCCCTGCCTCTTCTGGCCTGGTGTCTCCTCATGCCCCCTCAG 499

Qy     481 tgaggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     500 TGAGGATCTTCATGTACCTGCTCTTCTGTTTGCACACCCAGCATAGCCTCCTTGCAGGCA 559

Qy     541 gaaggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtcctgt 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     560 GAAGGCAGTAGGGCCCCTGCACACTCAGTTTCTCTCGTTTTCCTTAGTTATCAGTCCTGT 619

Qy     601 cctgtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggccct 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     620 CCTGTCCCACTCAGGTCTGTACTTAGGGCAGCTGGCCTGGATGGGCTTCACTGGGGCCCT 679

Qy     661 gtctgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgt 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     680 GTCTGTGTGCTGAGCCAGTTTCCCCTGCTGGCTGCAAGCTGTGGGTTCTTTCTCCTCTGT 739

Qy     721 gcccctcatgctgatcttctagatgccactcccaaatccccttcatacccaccaggatgt 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     740 GCCCCTCATGCTGATCTTCTAGATGCCACTCCCAAATCCCCTTCATACCCACCAGGATGT 799

Qy     781 gtgcccagccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggc 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     800 GTGCCCAGCCAGGCCTCCAGCACCCCCAGTGCAGCTCGTGATTGGAAACTCACCATCGGC 859

Qy     841 aggcagtggttcggtttaagagatggcattagagggagcccagtctggatgtggacttgg 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     860 AGGCAGTGGTTCGGTTTAAGAGATGGCATTAGAGGGAGCCCAGTCTGGATGTGGACTTGG 919

Qy     901 atgccctgtgggtatcagttctgctgacactttggcccgaaatagatccagtgctgagca 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     920 ATGCCCTGTGGGTATCAGTTCTGCTGACACTTTGGCCCGAAATAGATCCAGTGCTGAGCA 979

Qy     961 agcaatgtacaccggagcctcagtgagcccatctgcacagtggggagcatggagggatgg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     980 AGCAATGTACACCGGAGCCTCAGTGAGCCCATCTGCACAGTGGGGAGCATGGAGGGATGG 1039
```

```
Qy    1021 gtttggcctgtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtg 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1040 GTTTGGCCTGTGCTTCTGCTTATTCAGTCCTTCAGCTCACGGAAGGGATGCTAGTCCGTG 1099

Qy    1081 aaggtgacctcacagtactggttaattaaactttattgctcactgtcaaaaaaaaaaaaa 1140
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    1100 AAGGTGACCTCACAGTACTGGTTAATTAAACTTTATTGCTCACTGTTAAAAAAAAAAAAA 1159

Qy    1141 aaaaaaaaaaaa 1152
           ||||||||||||
Db    1160 AAAAAAAAAAAA 1171
```

```
RESULT   3
AB035343
LOCUS       AB035343     3144 bp    mRNA            PRI       12-DEC-2000
DEFINITION  Homo sapiens TCL6f4 mRNA for T-cell leukemia/lymphoma 6 ORF105,
            T-cell leukemia/lymphoma 6 ORF72, complete cds, clone:pDG4.
ACCESSION   AB035343
VERSION     AB035343.1  GI:8176590
KEYWORDS    T-cell leukemia/lymphoma 6 ORF72; TCL6f4; T-cell leukemia/lymphoma
            6 ORF105.
SOURCE      Homo sapiens cell_line:Daudi cDNA to mRNA, clone:pDG4.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3144)
  AUTHORS   Saitou,M., Sugimoto,J., Hatakeyama,T., Russo,G. and Isobe,M.
  TITLE     Identification of the TCL6 genes within the breakpoint cluster
            region on chromosome 14q32 in T-cell leukemia
  JOURNAL   Oncogene 19 (23), 2796-2802 (2000)
  MEDLINE   20309940
REFERENCE   2  (bases 1 to 3144)
  AUTHORS   Isobe,M., Saitou,M. and Sugimoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-NOV-1999) to the DDBJ/EMBL/GenBank databases.
            Masaharu Isobe, Toyama University, Materials and Biosystem
            Engineering, Faculty of Engineering; 3190 Gofuku, Toyama 930-8555,
            Japan (E-mail:isobe@eng.toyama-u.ac.jp., Tel:+81-76-445-6872,
            Fax:+81-76-445-6874)
FEATURES             Location/Qualifiers
     source          1. .3144
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /map="14q32.1"
                     /clone="pDG4"
                     /cell_line="Daudi"
     gene            1805. .3073
                     /gene="TCL6f4"
     CDS             1805. .2122
                     /gene="TCL6f4"
                     /codon_start=1
                     /product="T-cell leukemia/lymphoma 6 ORF105"
                     /protein_id="BAA96383.1"
                     /db_xref="GI:8176591"
                     /translation="MDAWGRLEARHKSYLHSRRGDHAKDKLQRDNRFASQTHSHWAKS
                     SCHCRFESRPFFLPSPSWSYNRSWGGESAERTWMKQRGTCLSHSPLDLLHPGLCQHKV
                     GAR"
     CDS             2855. .3073
                     /gene="TCL6f4"
                     /codon_start=1
                     /product="T-cell leukemia/lymphoma 6 ORF72"
                     /protein_id="BAA96384.1"
                     /db_xref="GI:8176592"
                     /translation="MALEGAQSGCGLGCPVGISSADTLARNRSSAEQAMYTGASVSPS
                     AQWGAWRDGFGLCFCLFSPSAHGRDASP"
BASE COUNT      764 a    809 c    809 g    762 t
ORIGIN
```

```
  Query Match            83.6%;  Score 963;  DB 9;  Length 3144;
```

```
  Best Local Similarity   100.0%;  Pred. No. 3e-255;
  Matches  963;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     190 tatgaacccagcatcacagtgcacttgtggcagatggcagtgcatacccgggagctactc 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2182 TATGAACCCAGCATCACAGTGCACTTGTGGCAGATGGCAGTGCATACCCGGGAGCTACTC 2241

Qy     250 tcctccggccagatgcccttctcccagctgcccgccgtgtggcagctctaccccgggagg 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2242 TCCTCCGGCCAGATGCCCTTCTCCCAGCTGCCCGCCGTGTGGCAGCTCTACCCCGGGAGG 2301

Qy     310 aagtaccgagcagcggattccagtttctgggaaatagcagaccatggccagattgactct 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2302 AAGTACCGAGCAGCGGATTCCAGTTTCTGGGAAATAGCAGACCATGGCCAGATTGACTCT 2361

Qy     370 atggagcagctggtcctaacatatcagccggagaggaaagactgacactgggagtggctg 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2362 ATGGAGCAGCTGGTCCTAACATATCAGCCGGAGAGGAAAGACTGACACTGGGAGTGGCTG 2421

Qy     430 gccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcagtgaggatct 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2422 GCCCTGCTGGCCCTGCCTCTTCTGGCCTGGTGTCTCCTCATGCCCCCTCAGTGAGGATCT 2481

Qy     490 tcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggcagaaggcagt 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2482 TCATGTACCTGCTCTTCTGTTTGCACACCCAGCATAGCCTCCTTGCAGGCAGAAGGCAGT 2541

Qy     550 agggcccctgcacactcagtttctctcgttttccttagttatcagtcctgtcctgtccca 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2542 AGGGCCCCTGCACACTCAGTTTCTCTCGTTTTCCTTAGTTATCAGTCCTGTCCTGTCCCA 2601

Qy     610 ctcaggtctgtacttagggcagctggcctggatgggcttcactggggccctgtctgtgtg 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2602 CTCAGGTCTGTACTTAGGGCAGCTGGCCTGGATGGGCTTCACTGGGGCCCTGTCTGTGTG 2661

Qy     670 ctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgtgcccctcat 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2662 CTGAGCCAGTTTCCCCTGCTGGCTGCAAGCTGTGGGTTCTTTCTCCTCTGTGCCCCTCAT 2721

Qy     730 gctgatcttctagatgccactcccaaatccccttcatacccaccaggatgtgtgcccagc 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2722 GCTGATCTTCTAGATGCCACTCCCAAATCCCCTTCATACCCACCAGGATGTGTGCCCAGC 2781

Qy     790 caggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggcaggcagtgg 849
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2782 CAGGCCTCCAGCACCCCCAGTGCAGCTCGTGATTGGAAACTCACCATCGGCAGGCAGTGG 2841

Qy     850 ttcggtttaagagatggcattagagggagcccagtctggatgtggacttggatgccctgt 909
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2842 TTCGGTTTAAGAGATGGCATTAGAGGGAGCCCAGTCTGGATGTGGACTTGGATGCCCTGT 2901

Qy     910 gggtatcagttctgctgacactttggcccgaaatagatccagtgctgagcaagcaatgta 969
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2902 GGGTATCAGTTCTGCTGACACTTTGGCCCGAAATAGATCCAGTGCTGAGCAAGCAATGTA 2961

Qy     970 caccggagcctcagtgagcccatctgcacagtggggagcatggagggatgggtttggcct 1029
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2962 CACCGGAGCCTCAGTGAGCCCATCTGCACAGTGGGGAGCATGGAGGGATGGGTTTGGCCT 3021

Qy    1030 gtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtgaaggtgacc 1089
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3022 GTGCTTCTGCTTATTCAGTCCTTCAGCTCACGGAAGGGATGCTAGTCCGTGAAGGTGACC 3081

Qy    1090 tcacagtactggttaattaaactttattgctcactgtcaaaaaaaaaaaaaaaaaaaaaa 1149
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3082 TCACAGTACTGGTTAATTAAACTTTATTGCTCACTGTCAAAAAAAAAAAAAAAAAAAAAA 3141

Qy    1150 aaa 1152
           |||
Db    3142 AAA 3144


RESULT   4
```

```
AB035342
LOCUS       AB035342     3240 bp    mRNA            PRI       12-DEC-2000
DEFINITION  Homo sapiens TCL6f3 mRNA for T-cell leukemia/lymphoma 6 ORF105,
            T-cell leukemia/lymphoma 6 ORF72, complete cds, clone:pDG3.
ACCESSION   AB035342
VERSION     AB035342.1  GI:8176587
KEYWORDS    TCL6f3; T-cell leukemia/lymphoma 6 ORF105; TML1 beta ORF72; TML1
            beta ORF105.
SOURCE      Homo sapiens cell_line:Daudi cDNA to mRNA, clone:pDG3.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3240)
  AUTHORS   Saitou,M., Sugimoto,J., Hatakeyama,T., Russo,G. and Isobe,M.
  TITLE     Identification of the TCL6 genes within the breakpoint cluster
            region on chromosome 14q32 in T-cell leukemia
  JOURNAL   Oncogene 19 (23), 2796-2802 (2000)
  MEDLINE   20309940
REFERENCE   2  (bases 1 to 3240)
  AUTHORS   Isobe,M., Saitou,M. and Sugimoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-NOV-1999) to the DDBJ/EMBL/GenBank databases.
            Masaharu Isobe, Toyama University, Materials and Biosystem
            Engineering, Faculty of Engineering; 3190 Gofuku, Toyama 930-8555,
            Japan (E-mail:isobe@eng.toyama-u.ac.jp., Tel:+81-76-445-6872,
            Fax:+81-76-445-6874)
FEATURES             Location/Qualifiers
     source          1. .3240
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /map="14q32.1"
                     /clone="pDG3"
                     /cell_line="Daudi"
     gene            1805. .3169
                     /gene="TCL6f3"
     CDS             1805. .2122
                     /gene="TCL6f3"
                     /codon_start=1
                     /product="T-cell leukemia/lymphoma 6 ORF105"
                     /protein_id="BAA96381.1"
                     /db_xref="GI:8176588"
                     /translation="MDAWGRLEARHKSYLHSRRGDHAKDKLQRDNRFASQTHSHWAKS
                     SCHCRFESRPFFLPSPSWSYNRSWGGESAERTWMKQRGTCLSHSPLDLLHPGLCQHKV
                     GAR"
     CDS             2951. .3169
                     /gene="TCL6f3"
                     /codon_start=1
                     /product="TML1 beta ORF72"
                     /protein_id="BAA96382.1"
                     /db_xref="GI:8176589"
                     /translation="MALEGAQSGCGLGCPVGISSADTLARNRSSAEQAMYTGASVSPS
                     AQWGAWRDGFGLCFCLFSPSAHGRDASP"
BASE COUNT      788 a    833 c    833 g    786 t
ORIGIN


  Query Match           83.6%;  Score 963;  DB 9;  Length 3240;
  Best Local Similarity   100.0%;  Pred. No. 3e-255;
  Matches  963;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     190 tatgaacccagcatcacagtgcacttgtggcagatggcagtgcatacccgggagctactc 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2278 TATGAACCCAGCATCACAGTGCACTTGTGGCAGATGGCAGTGCATACCCGGGAGCTACTC 2337

Qy     250 tcctccggccagatgcccttctcccagctgcccgccgtgtggcagctctaccccgggagg 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2338 TCCTCCGGCCAGATGCCCTTCTCCCAGCTGCCCGCCGTGTGGCAGCTCTACCCCGGGAGG 2397

Qy     310 aagtaccgagcagcggattccagtttctgggaaatagcagaccatggccagattgactct 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2398 AAGTACCGAGCAGCGGATTCCAGTTTCTGGGAAATAGCAGACCATGGCCAGATTGACTCT 2457
```

```
Qy     370 atggagcagctggtcctaacatatcagccggagaggaaagactgacactgggagtggctg 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2458 ATGGAGCAGCTGGTCCTAACATATCAGCCGGAGAGGAAAGACTGACACTGGGAGTGGCTG 2517

Qy     430 gccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcagtgaggatct 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2518 GCCCTGCTGGCCCTGCCTCTTCTGGCCTGGTGTCTCCTCATGCCCCCTCAGTGAGGATCT 2577

Qy     490 tcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggcagaaggcagt 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2578 TCATGTACCTGCTCTTCTGTTTGCACACCCAGCATAGCCTCCTTGCAGGCAGAAGGCAGT 2637

Qy     550 agggcccctgcacactcagtttctctcgttttccttagttatcagtcctgtcctgtccca 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2638 AGGGCCCCTGCACACTCAGTTTCTCTCGTTTTCCTTAGTTATCAGTCCTGTCCTGTCCCA 2697

Qy     610 ctcaggtctgtacttagggcagctggcctggatgggcttcactggggccctgtctgtgtg 669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2698 CTCAGGTCTGTACTTAGGGCAGCTGGCCTGGATGGGCTTCACTGGGGCCCTGTCTGTGTG 2757

Qy     670 ctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgtgcccctcat 729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2758 CTGAGCCAGTTTCCCCTGCTGGCTGCAAGCTGTGGGTTCTTTCTCCTCTGTGCCCCTCAT 2817

Qy     730 gctgatcttctagatgccactcccaaatccccttcatacccaccaggatgtgtgcccagc 789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2818 GCTGATCTTCTAGATGCCACTCCCAAATCCCCTTCATACCCACCAGGATGTGTGCCCAGC 2877

Qy     790 caggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggcaggcagtgg 849
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2878 CAGGCCTCCAGCACCCCCAGTGCAGCTCGTGATTGGAAACTCACCATCGGCAGGCAGTGG 2937

Qy     850 ttcggtttaagagatggcattagagggagcccagtctggatgtggacttggatgccctgt 909
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2938 TTCGGTTTAAGAGATGGCATTAGAGGGAGCCCAGTCTGGATGTGGACTTGGATGCCCTGT 2997

Qy     910 gggtatcagttctgctgacactttggcccgaaatagatccagtgctgagcaagcaatgta 969
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2998 GGGTATCAGTTCTGCTGACACTTTGGCCCGAAATAGATCCAGTGCTGAGCAAGCAATGTA 3057

Qy     970 caccggagcctcagtgagcccatctgcacagtggggagcatggagggatgggtttggcct 1029
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3058 CACCGGAGCCTCAGTGAGCCCATCTGCACAGTGGGGAGCATGGAGGGATGGGTTTGGCCT 3117

Qy    1030 gtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtgaaggtgacc 1089
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3118 GTGCTTCTGCTTATTCAGTCCTTCAGCTCACGGAAGGGATGCTAGTCCGTGAAGGTGACC 3177

Qy    1090 tcacagtactggttaattaaactttattgctcactgtcaaaaaaaaaaaaaaaaaaaaaa 1149
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3178 TCACAGTACTGGTTAATTAAACTTTATTGCTCACTGTCAAAAAAAAAAAAAAAAAAAAAA 3237

Qy    1150 aaa 1152
           |||
Db    3238 AAA 3240
```

```
RESULT   5
AB035340
LOCUS       AB035340     3532 bp    mRNA            PRI       12-DEC-2000
DEFINITION  Homo sapiens TCL6f1 mRNA for T-cell leukemia/lymphoma 6 ORF141,
            T-cell leukemia/lymphoma 6 ORF72, complete cds, clone:pDG1.
ACCESSION   AB035340
VERSION     AB035340.1  GI:8176581
KEYWORDS    TML1 beta ORF72; TCL6f1; T-cell leukemia/lymphoma 6 ORF72; T-cell
            leukemia/lymphoma 6 ORF141.
SOURCE      Homo sapiens cell_line:Daudi cDNA to mRNA, clone:pDG1.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3532)
  AUTHORS   Saitou,M., Sugimoto,J., Hatakeyama,T., Russo,G. and Isobe,M.
  TITLE     Identification of the TCL6 genes within the breakpoint cluster
            region on chromosome 14q32 in T-cell leukemia
  JOURNAL   Oncogene 19 (23), 2796-2802 (2000)
  MEDLINE   20309940
REFERENCE   2  (bases 1 to 3532)
  AUTHORS   Isobe,M., Saitou,M. and Sugimoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-NOV-1999) to the DDBJ/EMBL/GenBank databases.
            Masaharu Isobe, Toyama University, Materials and Biosystem
            Engineering, Faculty of Engineering; 3190 Gofuku, Toyama 930-8555,
            Japan (E-mail:isobe@eng.toyama-u.ac.jp., Tel:+81-76-445-6872,
            Fax:+81-76-445-6874)
FEATURES             Location/Qualifiers
     source          1. .3532
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /map="14q32.1"
                     /clone="pDG1"
                     /cell_line="Daudi"
     gene            1768. .3461
                     /gene="TCL6f1"
     CDS             1768. .2193
                     /gene="TCL6f1"
                     /codon_start=1
                     /product="T-cell leukemia/lymphoma 6 ORF141"
                     /protein_id="BAA96377.1"
                     /db_xref="GI:8176582"
                     /translation="MEPRVTQRKRPLDGCMGKITGITSDILKYDHKCFKLSLPAKFPE
                     VCGSDEVFPDPDLLHVLPVAGSLQQSIDQCCLQLESLCRPGLLCAHPTLLFKLHSSMK
                     NRPFFSLIYTYVKKTQQVRKRDRKPRGQVAAGPNPTSVM"
     CDS             3243. .3461
                     /gene="TCL6f1"
                     /codon_start=1
                     /product="T-cell leukemia/lymphoma 6 ORF72"
                     /protein_id="BAA96378.1"
                     /db_xref="GI:8176583"
                     /translation="MALEGAQSGCGLGCPVGISSADTLARNRSSAEQAMYTGASVSPS
                     AQWGAWRDGFGLCFCLFSPSAHGRDASP"
BASE COUNT      857 a    913 c    896 g    866 t
ORIGIN
```

```
  Query Match             83.6%;  Score 963;  DB 9;  Length 3532;
  Best Local Similarity   100.0%;  Pred. No. 3e-255;
  Matches  963;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     190 tatgaacccagcatcacagtgcacttgtggcagatggcagtgcatacccgggagctactc 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2570 TATGAACCCAGCATCACAGTGCACTTGTGGCAGATGGCAGTGCATACCCGGGAGCTACTC 2629

Qy     250 tcctccggccagatgcccttctcccagctgcccgccgtgtggcagctctaccccgggagg 309
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2630 TCCTCCGGCCAGATGCCCTTCTCCCAGCTGCCCGCCGTGTGGCAGCTCTACCCCGGGAGG 2689

Qy     310 aagtaccgagcagcggattccagtttctgggaaatagcagaccatggccagattgactct 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2690 AAGTACCGAGCAGCGGATTCCAGTTTCTGGGAAATAGCAGACCATGGCCAGATTGACTCT 2749

Qy     370 atggagcagctggtcctaacatatcagccggagaggaaagactgacactgggagtggctg 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2750 ATGGAGCAGCTGGTCCTAACATATCAGCCGGAGAGGAAAGACTGACACTGGGAGTGGCTG 2809

Qy     430 gccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcagtgaggatct 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2810 GCCCTGCTGGCCCTGCCTCTTCTGGCCTGGTGTCTCCTCATGCCCCCTCAGTGAGGATCT 2869

Qy     490 tcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggcagaaggcagt 549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2870 TCATGTACCTGCTCTTCTGTTTGCACACCCAGCATAGCCTCCTTGCAGGCAGAAGGCAGT 2929

Qy     550 agggcccctgcacactcagtttctctcgttttccttagttatcagtcctgtcctgtccca 609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2930 AGGGCCCCTGCACACTCAGTTTCTCTCGTTTTCCTTAGTTATCAGTCCTGTCCTGTCCCA 2989
```

```
Qy    610 ctcaggtctgtacttagggcagctggcctggatgggcttcactggggccctgtctgtgtg 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2990 CTCAGGTCTGTACTTAGGGCAGCTGGCCTGGATGGGCTTCACTGGGGCCCTGTCTGTGTG 3049

Qy    670 ctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgtgcccctcat 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3050 CTGAGCCAGTTTCCCCTGCTGGCTGCAAGCTGTGGGTTCTTTCTCCTCTGTGCCCCTCAT 3109

Qy    730 gctgatcttctagatgccactcccaaatccccttcatacccaccaggatgtgtgcccagc 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3110 GCTGATCTTCTAGATGCCACTCCCAAATCCCCTTCATACCCACCAGGATGTGTGCCCAGC 3169

Qy    790 caggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggcaggcagtgg 849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3170 CAGGCCTCCAGCACCCCCAGTGCAGCTCGTGATTGGAAACTCACCATCGGCAGGCAGTGG 3229

Qy    850 ttcggtttaagagatggcattagagggagcccagtctggatgtggacttggatgccctgt 909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3230 TTCGGTTTAAGAGATGGCATTAGAGGGAGCCCAGTCTGGATGTGGACTTGGATGCCCTGT 3289

Qy    910 gggtatcagttctgctgacactttggcccgaaatagatccagtgctgagcaagcaatgta 969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3290 GGGTATCAGTTCTGCTGACACTTTGGCCCGAAATAGATCCAGTGCTGAGCAAGCAATGTA 3349

Qy    970 caccggagcctcagtgagcccatctgcacagtggggagcatggagggatgggtttggcct 1029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3350 CACCGGAGCCTCAGTGAGCCCATCTGCACAGTGGGGAGCATGGAGGGATGGGTTTGGCCT 3409

Qy   1030 gtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtgaaggtgacc 1089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3410 GTGCTTCTGCTTATTCAGTCCTTCAGCTCACGGAAGGGATGCTAGTCCGTGAAGGTGACC 3469

Qy   1090 tcacagtactggttaattaaacttttattgctcactgtcaaaaaaaaaaaaaaaaaaaaaa 1149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3470 TCACAGTACTGGTTAATTAAACTTTATTGCTCACTGTCAAAAAAAAAAAAAAAAAAAAAA 3529

Qy   1150 aaa 1152
          |||
Db   3530 AAA 3532
```

```
RESULT   6
AB035341
LOCUS       AB035341     3599 bp    mRNA            PRI       12-DEC-2000
DEFINITION  Homo sapiens TCL6f2 mRNA for T-cell leukemia/lymphoma 6 ORF141,
            T-cell leukemia/lymphoma 6 ORF72, complete cds, clone:pDG2.
ACCESSION   AB035341
VERSION     AB035341.1  GI:8176584
KEYWORDS    T-cell leukemia/lymphoma 6 ORF72; TCL6f2; T-cell leukemia/lymphoma
            6 ORF141.
SOURCE      Homo sapiens cell_line:Daudi cDNA to mRNA, clone:pDG2.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3599)
  AUTHORS   Saitou,M., Sugimoto,J., Hatakeyama,T., Russo,G. and Isobe,M.
  TITLE     Identification of the TCL6 genes within the breakpoint cluster
            region on chromosome 14q32 in T-cell leukemia
  JOURNAL   Oncogene 19 (23), 2796-2802 (2000)
  MEDLINE   20309940
REFERENCE   2  (bases 1 to 3599)
  AUTHORS   Isobe,M., Saitou,M. and Sugimoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-NOV-1999) to the DDBJ/EMBL/GenBank databases.
            Masaharu Isobe, Toyama University, Materials and Biosystem
            Engineering, Faculty of Engineering; 3190 Gofuku, Toyama 930-8555,
            Japan (E-mail:isobe@eng.toyama-u.ac.jp., Tel:+81-76-445-6872,
            Fax:+81-76-445-6874)
FEATURES             Location/Qualifiers
     source          1. .3599
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /map="14q32.1"
                     /clone="pDG2"
                     /cell_line="Daudi"
     gene            1768. .3528
                     /gene="TCL6f2"
     CDS             1768. .2193
                     /gene="TCL6f2"
                     /codon_start=1
                     /product="T-cell leukemia/lymphoma 6 ORF141"
                     /protein_id="BAA96379.1"
                     /db_xref="GI:8176585"
                     /translation="MEPRVTQRKRPLDGCMGKITGITSDILKYDHKCFKLSLPAKFPE
                     VCGSDEVFPDPDLLHVLPVAGSLQQSIDQCCLQLESLCRPGLLCAHPTLLFKLHSSMK
                     NRPFFSLIYTYVKKTQQVRKRDRKPRGQVAAGPNPTSVM"
     CDS             3310. .3528
                     /gene="TCL6f2"
                     /codon_start=1
                     /product="T-cell leukemia/lymphoma 6 ORF72"
                     /protein_id="BAA96380.1"
                     /db_xref="GI:8176586"
                     /translation="MALEGAQSGCGLGCPVGISSADTLARNRSSAEQAMYTGASVSPS
                     AQWGAWRDGFGLCFCLFSPSAHGRDASP"
BASE COUNT      883 a    926 c    914 g    876 t
ORIGIN
```

Query Match 83.6%; Score 963; DB 9; Length 3599;
Best Local Similarity 100.0%; Pred. No. 3e-255;
Matches 963; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    190 tatgaacccagcatcacagtgcacttgtggcagatggcagtgcatacccgggagctactc 249
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2637 TATGAACCCAGCATCACAGTGCACTTGTGGCAGATGGCAGTGCATACCCGGGAGCTACTC 2696

Qy    250 tcctccggccagatgcccttctcccagctgcccgccgtgtggcagctctaccccgggagg 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2697 TCCTCCGGCCAGATGCCCTTCTCCCAGCTGCCCGCCGTGTGGCAGCTCTACCCCGGGAGG 2756

Qy    310 aagtaccgagcagcggattccagtttctgggaaatagcagaccatggccagattgactct 369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2757 AAGTACCGAGCAGCGGATTCCAGTTTCTGGGAAATAGCAGACCATGGCCAGATTGACTCT 2816

Qy    370 atggagcagctggtcctaacatatcagccggagaggaaagactgacactgggagtggctg 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2817 ATGGAGCAGCTGGTCCTAACATATCAGCCGGAGAGGAAAGACTGACACTGGGAGTGGCTG 2876

Qy    430 gccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcagtgaggatct 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2877 GCCCTGCTGGCCCTGCCTCTTCTGGCCTGGTGTCTCCTCATGCCCCCTCAGTGAGGATCT 2936

Qy    490 tcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggcagaaggcagt 549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2937 TCATGTACCTGCTCTTCTGTTTGCACACCCAGCATAGCCTCCTTGCAGGCAGAAGGCAGT 2996

Qy    550 agggcccctgcacactcagtttctctcgttttccttagttatcagtcctgtcctgtccca 609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2997 AGGGCCCCTGCACACTCAGTTTCTCTCGTTTTCCTTAGTTATCAGTCCTGTCCTGTCCCA 3056

Qy    610 ctcaggtctgtacttagggcagctggcctggatgggcttcactggggccctgtctgtgtg 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3057 CTCAGGTCTGTACTTAGGGCAGCTGGCCTGGATGGGCTTCACTGGGGCCCTGTCTGTGTG 3116

Qy    670 ctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgtgcccctcat 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3117 CTGAGCCAGTTTCCCCTGCTGGCTGCAAGCTGTGGGTTCTTTCTCCTCTGTGCCCCTCAT 3176

Qy    730 gctgatcttctagatgccactcccaaatccccttcatacccaccaggatgtgtgcccagc 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3177 GCTGATCTTCTAGATGCCACTCCCAAATCCCCTTCATACCCACCAGGATGTGTGCCCAGC 3236

Qy    790 caggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggcaggcagtgg 849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   3237 CAGGCCTCCAGCACCCCCAGTGCAGCTCGTGATTGGAAACTCACCATCGGCAGGCAGTGG 3296

Qy    850 ttcggtttaagagatggcattagagggagcccagtctggatgtggacttggatgccctgt 909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3297 TTCGGTTTAAGAGATGGCATTAGAGGGAGCCCAGTCTGGATGTGGACTTGGATGCCCTGT 3356

Qy    910 gggtatcagttctgctgacactttggcccgaaatagatccagtgctgagcaagcaatgta 969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3357 GGGTATCAGTTCTGCTGACACTTTGGCCCGAAATAGATCCAGTGCTGAGCAAGCAATGTA 3416

Qy    970 caccggagcctcagtgagcccatctgcacagtggggagcatggagggatgggtttggcct 1029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3417 CACCGGAGCCTCAGTGAGCCCATCTGCACAGTGGGGAGCATGGAGGGATGGGTTTGGCCT 3476

Qy   1030 gtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtgaaggtgacc 1089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3477 GTGCTTCTGCTTATTCAGTCCTTCAGCTCACGGAAGGGATGCTAGTCCGTGAAGGTGACC 3536

Qy   1090 tcacagtactggttaattaaactttattgctcactgtcaaaaaaaaaaaaaaaaaaaaaa 1149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3537 TCACAGTACTGGTTAATTAAACTTTATTGCTCACTGTCAAAAAAAAAAAAAAAAAAAAAA 3596

Qy   1150 aaa 1152
          |||
Db   3597 AAA 3599


RESULT   7
AC010356/c
LOCUS       AC010356   114438 bp    DNA             HTG       05-MAY-2000
DEFINITION  Homo sapiens chromosome 5 clone CTD-2028C21, WORKING DRAFT
            SEQUENCE, 10 unordered pieces.
ACCESSION   AC010356
VERSION     AC010356.4  GI:7710750
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 114438)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 114438)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On May 5, 2000 this sequence version replaced gi:7341483.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 638370
            Center clone name: CITB-H1_2028C21
            -------
            Summary Statistics
            Consensus quality: 105028 bases at least Q40
            Consensus quality: 110839 bases at least Q30
            Consensus quality: 111997 bases at least Q20
            Estimated insert size: 123000; pulse field gel estimation
            Estimated insert size: 113538; sum-of-contigs estimation
            Quality coverage:  4.61 in Q20 bases; pulse field gel estimation
            Quality coverage:  4.99 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2455: contig of 2455 bp in length
            *     2456     2555: gap of unknown length
            *     2556     5533: contig of 2978 bp in length
            *     5534     5633: gap of unknown length
            *     5634     9013: contig of 3380 bp in length
            *     9014     9113: gap of unknown length
            *     9114    13014: contig of 3901 bp in length
            *    13015    13114: gap of unknown length
            *    13115    19656: contig of 6542 bp in length
            *    19657    19756: gap of unknown length
            *    19757    25844: contig of 6088 bp in length
            *    25845    25944: gap of unknown length
            *    25945    37665: contig of 11721 bp in length
            *    37666    37765: gap of unknown length
            *    37766    61723: contig of 23958 bp in length
            *    61724    61823: gap of unknown length
            *    61824    86508: contig of 24685 bp in length
            *    86509    86608: gap of unknown length
            *    86609   114438: contig of 27830 bp in length.
FEATURES             Location/Qualifiers
     source          1. .114438
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2028C21"
                     /clone_lib="CalTech human BAC library D"
BASE COUNT    36135 a  22103 c  21741 g  33555 t    904 others
ORIGIN


  Query Match             62.3%;  Score 718;  DB 2;  Length 114438;
  Best Local Similarity   81.9%;  Pred. No. 1.5e-187;
  Matches  945;  Conservative    0;  Mismatches  155;  Indels   54;  Gaps    8;

Qy      2 aggcgggtcccggttgcagacttgccatggcctccgaagcttctgtgcgtctaggggtgc 61
          |||||   || |  ||||| ||||||||||||||| ||||||||  || |||  |||  |
Db  74977 AGGCGACCCCAGAGTGCAGCCTTGCCATGGCCTCCAAAGCTTCTCCGCATCTCAGGGCAC 74918

Qy     62 cccctgg--ccgtctgtggatccagaggcctggcatctacgaagatgaggaggggagaac 119
          ||||     | ||||||||||  ||| |||||||||| || |||||  ||||||| ||
Db  74917 CCCCCCCAACTGTCTGTGGATCTGGAGACCTGGCATCTATGAGGATGACAAGGGGAGGAC 74858

Qy    120 ctgggtgactgtggtcgtgcggttcaatccctcgcgtagggaatgggccagg------gc 173
          |||||||||||||||  |||||  ||| ||||  ||  || ||| ||||||||
Db  74857 CTGGGTGACTGTGGTTGTGCGCCTCAGTCCCCTGCACAGAGAAAGGGCCAGGTGCTCAAG 74798

Qy    174 ctcccagggcagcagatatgaacccagcatcacagtgcacttgtggcagatggcagtgca 233
          |||||||||||||| |||||||| |||||||||||||| ||||| ||||| |||||||
Db  74797 CTCCCAGGGCAGCACATATGAACCTAGCATCACAGTGCACATGTGGTAGATGCCAGTGCA 74738

Qy    234 tacccgggagctactctcctccggccagatgcccttctcccagctgcccgccgtgtggca 293
           |||||||| | ||| ||| || ||||| |||||||||||||| ||||   ||||||||||
Db  74737 CACCCGGGAACCACTGTCCCCCAGCCAGCTGCCCTTCTCCCACCTGCTTCCCGTGTGGCA 74678

Qy    294 gctctaccccgggaggaagtaccgagcagcggattccagtttctgggaaatagcagacca 353
          ||||||||||  ||||| ||| | ||||||  ||||||||||||||||||||||  ||
Db  74677 GCTCTACCCCAAGAGGAGGTATCAAGCAGCAAATTCCAGTTTCTGGGAAATAGTGGA--- 74619

Qy    354 tggccagattgactctatggagcagctggtcctaacatatcagccggagaggaaagactg 413
             |||||| | ||| |||||||||||||||||||||||  |||||  |||||  ||||
Db  74620 ---CCAGATCGTCTCCATGGAGCAGCTGGTCCTAACATATTGGCCGGCAAGGAATAACTG 74564

Qy    414 acactgggagtggctggccctgctggccctgcctcttctggcctggtgtctcctcatgcc 473
          |                              ||| ||||||||  |||||| ||  |||
Db  74563 A------------------------------CTCCTCTGGCCTCATGTCTCTTCGGGCC 74535

Qy    474 ccctcagtgaggatcttcatgtacctgctcttctgtttgcacacccagcatagcctcctt 533
          ||||||||||||||||  ||||| |||| ||| |||||||||||||| | ||||||||
Db  74534 CCCTCAGTGAGGATCTTTGTGTACTTGCTATTCCGTTTGCACACCCAGCGTGGCCTCCTT 74475

Qy    534 gcaggcagaaggcagtagggcccctgcacactcagtttctctcgttttccttagttatca 593
          |||||||| ||||||| |||||||||| ||||||| ||||||| ||||| | | ||||||
```

```
Db  74474 GCAGGCAGGAGGCAGTGGGGCCCCTGCCCACTCAGCTTCTCTCATTTTCTTCACTTATCA 74415

Qy    594 gtcctgtcctgtcccactcaggtctgtacttagggcagctggcctggatgggcttcact- 652
          ||| |||||||| |||||||| |||  ||| ||||||| ||||||||||||||||||||
Db  74414 GTCTTGTCCTGTTCCACTCAGATCTACACTGAGGGCAGTTGGCCTGGATGGGCTTCACTA 74355

Qy    653 ggggccctgtctgtgtgctgagccagtttcccctgctggctgcaagctgtgggttc--tt 710
          ||||||| |||||||  ||| ||| |||||||||||||||||||||||  |||||||  ||
Db  74354 GGGGCCCCGTCTGTGCACTGGGCCCGTTTCCCCTGCTGGCTGCAAGCCATGGGTTCTTTT 74295

Qy    711 tctcctctgtgcccctcatgctgatcttctagatgccactcccaaatcccctt-----ca 765
          |||||||| |||||||||||||| ||||||||||||||||||||||||||||     ||
Db  74294 TCTCCTCTCTGCCCCTCATGCTGACCTTCTAGATGCCACTCCCAAATCCCCTTCACTCCA 74235

Qy    766 tacccaccaggatgtgtgcccagccaggcctccagcacccccagtgcagctcgtgattgg 825
          |||||||||| |   |||||| |||||||||  |||||| ||||||||| | |||||||
Db  74234 TACCCACCAGGCTTCATGCCCACCCAGGCCTCTGGCACCCTCAGTGCAGCCCATGATTGG 74175

Qy    826 aaactcaccatcggcaggcagtggttcggtttaagagatggcattagagggagcccagtc 885
          |||||||||||| |||| |||||| |||||||||||||| |||   ||||||| |   |||
Db  74174 GAACTCACCATCAGCAGTCAGTGGCTCGGTTTAAGAGAGGGCCGCAGAGGGAACTGGGTC 74115

Qy    886 tggatgtggacttggatgccctgtgggtatcagttctgctgacactttggcccgaaatag 945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  74114 TGGATGTGGACTTGGATGCCCTGTGGGTATCAGTTCTGCTGACACTTTGGCCTGAAATAG 74055

Qy    946 atccagtgctgagcaagcaatgtacaccggagcctcagtgagcccatctgcacagtgggg 1005
          ||  |||||||||||||||| |||| | |||||||||| || ||||||||||||||||||
Db  74054 ATAAAGTGCTGAGCAAGCAGTGTATGCTGGAGCCTCAGTAAGTCCATCTGCACAGTGGGG 73995

Qy   1006 agcatggagggatgggtttggcctgtgcttctgcttattcagtccttcagctcacggaag 1065
          ||| |||||||||||||||| | |||||||||| ||||||| |||||||||| | |
Db  73994 AGCGTGGAGGGATGGGTTTGGCTTATGCTTCTGCTTGTTCAGTCTTTCAGCTCACAGCA- 73934

Qy   1066 ggatgctagtccgtgaaggtgacctcacagtactggttaattaaactttattgctcactg 1125
          |||||||| ||| |||||||||||||| |||||||||||||||| ||||||||||||||
Db  73935 GGATGCTAATCCATGAAGGTGACCTCGCAGTACTGGTTAATTAAGCTTTATTGCTCACTG 73876

Qy   1126 tcaaaaaaaaaaaa 1139
             ||| ||||||||
Db  73875 AAAAAGAAAAAAAA 73862


RESULT   8
AC010359/c
LOCUS       AC010359   204843 bp    DNA             PRI       22-DEC-2000
DEFINITION  Homo sapiens chromosome 5 clone CTD-2033C11, complete sequence.
ACCESSION   AC010359
VERSION     AC010359.5  GI:11968292
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 204843)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 204843)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 204843)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-DEC-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Dec 22, 2000 this sequence version replaced gi:7710748.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.7% of Sequence;
            Estimated Total Number of Errors is 0.5.
            STS Content:
            WI-31186 G24491.
FEATURES             Location/Qualifiers
     source          1. .204843
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2033C11"
BASE COUNT    66824 a  40891 c  37481 g  59647 t
ORIGIN


  Query Match           62.3%;  Score 718;  DB 9;  Length 204843;
  Best Local Similarity   81.9%;  Pred. No. 1.5e-187;
  Matches  945;  Conservative    0;  Mismatches  155;  Indels   54;  Gaps    8;

Qy      2 aggcgggtcccggttgcagacttgccatggcctccgaagcttctgtgcgtctaggggtgc 61
          |||||   || |  ||||| ||||||||||||||| ||||||||  || |||  |||  |
Db  95870 AGGCGACCCCAGAGTGCAGCCTTGCCATGGCCTCCAAAGCTTCTCCGCATCTCAGGGCAC 95811

Qy     62 cccctgg--ccgtctgtggatccagaggcctggcatctacgaagatgaggaggggagaac 119
          ||||     | |||||||||||  ||| |||||||||| || |||||  ||||||| ||
Db  95810 CCCCCCCAACTGTCTGTGGATCTGGAGACCTGGCATCTATGAGGATGACAAGGGGAGGAC 95751

Qy    120 ctgggtgactgtggtcgtgcggttcaatccctcgcgtagggaatgggccagg------gc 173
          |||||||||||||| |||||  ||| ||||  ||  || ||| ||||||||
Db  95750 CTGGGTGACTGTGGTTGTGCGCCTCAGTCCCCTGCACAGAGAAAGGGCCAGGTGCTCAAG 95691

Qy    174 ctcccagggcagcagatatgaacccagcatcacagtgcacttgtggcagatggcagtgca 233
          |||||||||||||| |||||||| |||||||||||||| ||||| ||||| |||||||
Db  95690 CTCCCAGGGCAGCACATATGAACCTAGCATCACAGTGCACATGTGGTAGATGCCAGTGCA 95631

Qy    234 tacccgggagctactctcctccggccagatgcccttctcccagctgcccgccgtgtggca 293
          ||||||||  ||| ||| || ||||| |||||||||||||| ||||   ||||||||||
Db  95630 CACCCGGGAACCACTGTCCCCCAGCCAGCTGCCCTTCTCCCACCTGCTTCCCGTGTGGCA 95571

Qy    294 gctctaccccgggaggaagtaccgagcagcggattccagtttctgggaaatagcagacca 353
          ||||||||||  |||| ||| | ||||||  |||||||||||||||||||||  ||
Db  95570 GCTCTACCCCAAGAGGAGGTATCAAGCAGCAAATTCCAGTTTCTGGGAAATAGTGGA--- 95512

Qy    354 tggccagattgactctatggagcagctggtcctaacatatcagccggagaggaaagactg 413
          ||||||  | ||| ||||||||||||||||||||||||  |||||  |||||  ||||
Db  95513 ---CCAGATCGTCTCCATGGAGCAGCTGGTCCTAACATATTGGCCGGCAAGGAATAACTG 95457

Qy    414 acactgggagtggctggccctgctggccctgcctcttctggcctggtgtctcctcatgcc 473
          |                              ||| |||||||  |||||| ||  |||
Db  95456 A------------------------------CTCCTCTGGCCTCATGTCTCTTCGGGCC 95428

Qy    474 ccctcagtgaggatcttcatgtacctgctcttctgtttgcacacccagcatagcctcctt 533
          ||||||||||||||||  |||||  |||| |||  |||||||||||||| | ||||||||
Db  95427 CCCTCAGTGAGGATCTTTGTGTACTTGCTATTCCGTTTGCACACCCAGCGTGGCCTCCTT 95368

Qy    534 gcaggcagaaggcagtagggcccctgcacactcagtttctctcgttttccttagttatca 593
          |||||||| ||||||| |||||||||| ||||||| ||||||| |||||  |  | ||||||
Db  95367 GCAGGCAGGAGGCAGTGGGGCCCCTGCCCACTCAGCTTCTCTCATTTTCTTCACTTATCA 95308

Qy    594 gtcctgtcctgtcccactcaggtctgtacttagggcagctggcctggatgggcttcact- 652
          ||| |||||||| |||||||| |||  ||| ||||||| ||||||||||||||||||||
Db  95307 GTCTTGTCCTGTTCCACTCAGATCTACACTGAGGGCAGTTGGCCTGGATGGGCTTCACTA 95248

Qy    653 ggggccctgtctgtgtgctgagccagtttcccctgctggctgcaagctgtgggttc--tt 710
          ||||||| |||||||  ||| ||| |||||||||||||||||||||||  |||||||  ||
Db  95247 GGGGCCCCGTCTGTGCACTGGGCCCGTTTCCCCTGCTGGCTGCAAGCCATGGGTTCTTTT 95188

Qy    711 tctcctctgtgcccctcatgctgatcttctagatgccactcccaaatcccctt-----ca 765
          |||||||| |||||||||||||| ||||||||||||||||||||||||||||     ||
Db  95187 TCTCCTCTCTGCCCCTCATGCTGACCTTCTAGATGCCACTCCCAAATCCCCTTCACTCCA 95128

Qy    766 tacccaccaggatgtgtgcccagccaggcctccagcacccccagtgcagctcgtgattgg 825
          |||||||||| |   |||||| |||||||||  |||||| ||||||||| | |||||||
Db  95127 TACCCACCAGGCTTCATGCCCACCCAGGCCTCTGGCACCCTCAGTGCAGCCCATGATTGG 95068
```

```
Qy    826 aaactcaccatcggcaggcagtggttcggtttaagagatggcattagagggagcccagtc 885
          ||||||||||| |||| |||||| |||||||||||| |||   ||||||| |   |||
Db  95067 GAACTCACCATCAGCAGTCAGTGGCTCGGTTTAAGAGAGGGCCGCAGAGGGAACTGGGTC 95008

Qy    886 tggatgtggacttggatgccctgtgggtatcagttctgctgacactttggcccgaaatag 945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  95007 TGGATGTGGACTTGGATGCCCTGTGGGTATCAGTTCTGCTGACACTTTGGCCTGAAATAG 94948

Qy    946 atccagtgctgagcaagcaatgtacaccggagcctcagtgagcccatctgcacagtgggg 1005
          ||  |||||||||||||| |||| | |||||||||| || ||||||||||||||||||
Db  94947 ATAAAGTGCTGAGCAAGCAGTGTATGCTGGAGCCTCAGTAAGTCCATCTGCACAGTGGGG 94888

Qy   1006 agcatggagggatgggtttggcctgtgcttctgcttattcagtccttcagctcacggaag 1065
          ||| |||||||||||||||| | |||||||||| ||||||| |||||||||| | |
Db  94887 AGCGTGGAGGGATGGGTTTGGCTTATGCTTCTGCTTGTTCAGTCTTTCAGCTCACAGCA- 94827

Qy   1066 ggatgctagtccgtgaaggtgacctcacagtactggttaattaaactttattgctcactg 1125
          |||||||| ||| |||||||||||||| |||||||||||||||| ||||||||||||||
Db  94828 GGATGCTAATCCATGAAGGTGACCTCGCAGTACTGGTTAATTAAGCTTTATTGCTCACTG 94769

Qy   1126 tcaaaaaaaaaaaa 1139
            ||| ||||||||
Db  94768 AAAAAGAAAAAAAA 94755
```

```
RESULT   9
AC008889
LOCUS       AC008889   137808 bp    DNA             HTG       19-APR-2001
DEFINITION  Homo sapiens chromosome 5 clone CTD-2221G1, WORKING DRAFT SEQUENCE,
            12 ordered pieces.
ACCESSION   AC008889
VERSION     AC008889.4  GI:7711367
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 137808)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 137808)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On May 6, 2000 this sequence version replaced gi:7709381.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 712558
            Center clone name: CITB-H1_2221G1
            -------
            Summary Statistics
            Consensus quality: 129558 bases at least Q40
            Consensus quality: 135125 bases at least Q30
            Consensus quality: 136017 bases at least Q20
            Estimated insert size: 143000; pulse field gel estimation
            Estimated insert size: 137258; sum-of-contigs estimation
            Quality coverage:  7.91 in Q20 bases; pulse field gel estimation
            Quality coverage:  8.24 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1     7204: contig of 7204 bp in length
            *     7205     7304: gap of unknown length
            *     7305    31808: contig of 24504 bp in length
            *    31809    31908: gap of unknown length
            *    31909    34114: contig of 2206 bp in length
            *    34115    34214: gap of unknown length
            *    34215    40106: contig of 5892 bp in length
            *    40107    40206: gap of unknown length
            *    40207    49908: contig of 9702 bp in length
            *    49909    50008: gap of unknown length
            *    50009    69581: contig of 19573 bp in length
            *    69582    69681: gap of unknown length
            *    69682    91290: contig of 21609 bp in length
            *    91291    91390: gap of unknown length
            *    91391   116545: contig of 25155 bp in length
            *   116546   116645: gap of unknown length
            *   116646   118981: contig of 2336 bp in length
            *   118982   119081: gap of unknown length
            *   119082   132967: contig of 13886 bp in length
            *   132968   133067: gap of unknown length
            *   133068   134154: contig of 1087 bp in length
            *   134155   134254: gap of unknown length
            *   134255   137808: contig of 3554 bp in length.
FEATURES             Location/Qualifiers
     source          1. .137808
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2221G1"
                     /clone_lib="CalTech human BAC library D"
BASE COUNT    39509 a  25657 c  28235 g  43290 t   1117 others
ORIGIN
```

```
  Query Match             60.9%;  Score 701.6;  DB 2;  Length 137808;
  Best Local Similarity   81.0%;  Pred. No. 5.3e-183;
  Matches  935;  Conservative    0;  Mismatches  165;  Indels   55;  Gaps    8;

Qy      2 aggcgggtcccggttgcagacttgccatggcctccgaagcttctgtgcgtctaggggtgc 61
          |||||   || |  |||||  |||||||||||||||| ||||||||  || |||  |||  |
Db 104445 AGGCGACCCCAGAGTGCAGCCTTGCCATGGCCTCCAAAGCTTCTCCGCATCTCAGGGCAC 104504

Qy     62 cccctgg--ccgtctgtggatccagaggcctggcatctacgaagatgaggaggggagaac 119
          ||||     | ||||||||||||  ||| |||||||||||| || |||||  ||||||| ||
Db 104505 CCCCCCCAACTGTCTGTGGATCTGGAGACCTGGCATCTATGAGGATGACAAGGGGAGGAC 104564

Qy    120 ctgggtgactgtggtcgtgcggttcaatccctcgcgtagggaatgggccagg-------g 172
          |||||||||||||||| |||||  ||| ||||  ||  || ||| ||||||||       |
Db 104565 CTGGGTGACTGTGGTTGTGCGCCTCAGTCCCCTGCACAGAGAAAGGGCCAGGTGCTCAAG 104624

Qy    173 cctcccagggcagcagatatgaacccagcatcacagtgcacttgtggcagatggcagtgc 232
          |  ||  |||||||| ||||||||| |||||||||||||||| ||||| ||||| ||||||
Db 104625 CTCCCAGGGGCAGCACATATGAACCTAGCATCACAGTGCACATGTGGTAGATGCCAGTGC 104684

Qy    233 atacccgggagctactctcctccggccagatgcccttctcccagctgcccgccgtgtggc 292
          | |||||||| | ||| ||| || ||||| ||||||||||||| ||||   |||||||||
Db 104685 ACACCCGGGAACCACTGTCCCCCAGCCAGCTGCCCTTCTCCCACCTGCTTCCCGTGTGGC 104744

Qy    293 agctctaccccgggaggaagtaccgagcagcggattccagtttctgggaaatagcagacc 352
          ||||||||||| ||||| ||| | ||||||  ||||||||||||||||||||||  ||
Db 104745 AGCTCTACCCCAAGAGGAGGTATCAAGCAGCAAATTCCAGTTTCTGGGAAATAGTGGA-- 104802

Qy    353 atggccagattgactctatggagcagctggtcctaacatatcagccggagaggaaagact 412
            |||||| | ||| ||||||||||||||||||||||||  |||||  |||||  |||
Db 104803 ----CCAGATCGTCTCCATGGAGCAGCTGGTCCTAACATATTGGCCGGCAAGGAATAACT 104858

Qy    413 gacactgggagtggctggccctgctggccctgcctcttctggcctggtgtctcctcatgc 472
          ||                                 ||| |||||||| |||||| ||  ||
Db 104859 GA------------------------------CTCCTCTGGCCTCATGTCTCTTCGGGC 104887

Qy    473 cccctcagtgaggatcttcatgtacctgctcttctgtttgcacacccagcatagcctcct 532
          |||||||||||||||||  ||||| |||| ||| ||||||||||||||| | |||||||
```

```
Db 104888 CCCCTCAGTGAGGATCTTTGTGTACTTGCTATTCCGTTTGCACACCCAGCGTGGCCTCCT 104947

Qy    533 tgcaggcagaaggcagtagggcccctgcacactcagtttctctcgttttccttagttatc 592
          ||||||||| ||||||| |||||||||| ||||||| ||||||| ||||| | | |||||
Db 104948 TGCAGGCAGGAGGCAGTGGGGCCCCTGCCCACTCAGCTTCTCTCATTTTCTTCACTTATC 105007

Qy    593 agtcctgtcctgtcccactcaggtctgtacttagggcagctggcctggatgggcttcact 652
          |||| |||||||| |||||||| |||  ||| ||||||| ||||||||||||||||||||
Db 105008 AGTCTTGTCCTGTTCCACTCAGATCTACACTGAGGGCAGTTGGCCTGGATGGGCTTCACT 105067

Qy    653 -ggggccctgtctgtgtgctgagccagtttcccctgctggctgcaagctgtgggttc--t 709
           ||||||| |||||||  ||| ||| ||||||||||||||||||||||  |||||||  |
Db 105068 AGGGGCCCCGTCTGTGCACTGGGCCCGTTTCCCCTGCTGGCTGCAAGCCATGGGTTCTTT 105127

Qy    710 ttctcctctgtgcccctcatgctgatcttctagatgccactcccaaatcccctt-----c 764
          ||||||||| |||||||||||||| ||||||||||||||||||||||||||     |
Db 105128 TTCTCCTCTCTGCCCCTCATGCTGACCTTCTAGATGCCACTCCCAAATCCCCTTCACTCC 105187

Qy    765 atacccaccaggatgtgtgcccagccaggcctccagcacccccagtgcagctcgtgattg 824
          |||||||||||| |   |||||| |||||||||  |||||| ||||||||| | ||||||
Db 105188 ATACCCACCAGGCTTCATGCCCACCCAGGCCTCTGGCACCCTCAGTGCAGCCCATGATTG 105247

Qy    825 gaaactcaccatcggcaggcagtggttcggtttaagagatggcattagagggagcccagt 884
          | |||||||||| |||| |||||| |||||||||||| |||   ||||||| |   ||
Db 105248 GGAACTCACCATCAGCAGTCAGTGGCTCGGTTTAAGAGAGGGCCGCAGAGGGAACTGGGT 105307

Qy    885 ctggatgtggacttggatgccctgtgggtatcagttctgctgacactttggcccgaaata 944
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db 105308 CTGGATGTGGACTTGGATGCCCTGTGGGTATCAGTTCTGCTGACACTTTGGCCTGAAATA 105367

Qy    945 gatccagtgctgagcaagcaatgtacaccggagcctcagtgagcccatctgcacagtggg 1004
          |||  |||||||||| |||| |||| | |||||||||| || ||||||||||||||||
Db 105368 GATAAAGTGCTGAGC-AGCAGTGTATGCTGGAGCCTCAGTAAGTCCATCTGCACAGTGGG 105426

Qy   1005 gagcatggagggatgggtttggcctgtgcttctgcttattcagtccttcagctcacggaa 1064
          |||| |||||||||||||||| | |||||||||| ||||||| |||||||||| | |
Db 105427 GAGCGTGGAGGGATGGGTTTGGCTTATGCTTCTGCTTGTTCAGTCTTTCAGCTCACAGCA 105486

Qy   1065 gggatgctagtccgtgaaggtgacctcacagtactggttaattaaactttattgctcact 1124
          ||      |   | ||||||||||||| |||||||||||||||| ||||||||||||||
Db 105487 GGATGCTAATCNCATGAAGGTGACCTCGCAGTACTGGTTAATTAAGCTTTATTGCTCACT 105546

Qy   1125 gtcaaaaaaaaaaaa 1139
          |  ||| ||||||||
Db 105547 GAAAAAGAAAAAAAA 105561


RESULT  10
AF110465
LOCUS       AF110465     6486 bp    DNA             PRI       18-MAR-1999
DEFINITION  Homo sapiens T-cell leukemia/lymphoma 1B (TCL1B) gene, complete
            cds.
ACCESSION   AF110465
VERSION     AF110465.1  GI:4324702
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6486)
  AUTHORS   Pekarsky,Y., Hallas,C., Isobe,M., Russo,G. and Croce,C.M.
  TITLE     Abnormalities at 14q32.1 in T cell malignancies involve two
            oncogenes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2949-2951 (1999)
  MEDLINE   99178995
REFERENCE   2  (bases 1 to 6486)
  AUTHORS   Pekarsky,Y., Hallas,C. and Croce,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-1998) KCI, Thomas Jefferson University, 233 South
            10th Str. BLSB 1032, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .6486
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /map="14q32.1"
     mRNA            join(141. .329,4439. .4609,4971. .5039,5634. .6330)
                     /gene="TCL1B"
     gene            141. .6330
                     /gene="TCL1B"
     CDS             join(168. .329,4439. .4609,4971. .5024)
                     /gene="TCL1B"
                     /codon_start=1
                     /product="T-cell leukemia/lymphoma 1B"
                     /protein_id="AAD16996.1"
                     /db_xref="GI:4324703"
                     /translation="MASEASVRLGVPPGRLWIQRPGIYEDEEGRTWVTVVVRFNPSRR
                     EWARASQGSRYEPSITVHLWQMAVHTRELLSSGQMPFSQLPAVWQLYPGRKYRAADSS
                     FWEIADHGQIDSMEQLVLTYQPERKD"
BASE COUNT     1600 a   1559 c   1687 g   1639 t      1 others
ORIGIN


  Query Match             60.9%;  Score 701.2;  DB 9;  Length 6486;
  Best Local Similarity   99.6%;  Pred. No. 7.2e-183;
  Matches  703;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy    424 tggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcagtga 483
          | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5628 TCGCAGGCCCTGCTGGCCCTGCCTCTTCTGGCCTGGTGTCTCCTCATGCCCCCTCAGTGA 5687

Qy    484 ggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggcagaa 543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5688 GGATCTTCATGTACCTGCTCTTCTGTTTGCACACCCAGCATAGCCTCCTTGCAGGCAGAA 5747

Qy    544 ggcagtagggccccctgcacactcagtttctctcgttttccttagttatcagtcctgtcct 603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5748 GGCAGTAGGGCCCCTGCACACTCAGTTTCTCTCGTTTTCCTTAGTTATCAGTCCTGTCCT 5807

Qy    604 gtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggccctgtc 663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5808 GTCCCACTCAGGTCTGTACTTAGGGCAGCTGGCCTGGATGGGCTTCACTGGGGCCCTGTC 5867

Qy    664 tgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgtgcc 723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5868 TGTGTGCTGAGCCAGTTTCCCCTGCTGGCTGCAAGCTGTGGGTTCTTTCTCCTCTGTGCC 5927

Qy    724 cctcatgctgatcttctagatgccactcccaaatcccttcatacccaccaggatgtgtg 783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5928 CCTCATGCTGATCTTCTAGATGCCACTCCCAAATCCCCTTCATACCCACCAGGATGTGTG 5987

Qy    784 cccagccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggcagg 843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5988 CCCAGCCAGGCCTCCAGCACCCCCAGTGCAGCTCGTGATTGGAAACTCACCATCGGCAGG 6047

Qy    844 cagtggttcggtttaagagatggcattagagggagcccagtctggatgtggacttggatg 903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6048 CAGTGGTTCGGTTTAAGAGATGGCATTAGAGGGAGCCCAGTCTGGATGTGGACTTGGATG 6107

Qy    904 ccctgtgggtatcagttctgctgacactttggcccgaaatagatccagtgctgagcaagc 963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6108 CCCTGTGGGTATCAGTTCTGCTGACACTTTGGCCCGAAATAGATCCAGTGCTGAGCAAGC 6167

Qy    964 aatgtacaccggagcctcagtgagcccatctgcacagtggggagcatggagggatgggtt 1023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6168 AATGTACACCGGAGCCTCAGTGAGCCCATCTGCACAGTGGGGAGCATGGAGGGATGGGTT 6227

Qy   1024 tggcctgtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtgaag 1083
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6228 TGGCCTGTGCTTCTGCTTATTCAGTCCTTCAGCTCACGGAAGGGATGCTAGTCCGTGAAG 6287

Qy   1084 gtgacctcacagtactggttaattaaactttattgctcactgtcaa 1129
          |||||||||||||||||||||||||||||||||||||||||||| |
Db   6288 GTGACCTCACAGTACTGGTTAATTAAACTTTATTGCTCACTGTCCA 6333
```

```
RESULT  11
AF110467
LOCUS       AF110467     1248 bp    DNA             PRI       18-MAR-1999
DEFINITION  Homo sapiens T-cell leukemia/lymphoma 1B (TCL1B) pseudogene,
            complete sequence.
ACCESSION   AF110467
VERSION     AF110467.1  GI:4324706
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1248)
  AUTHORS   Pekarsky,Y., Hallas,C., Isobe,M., Russo,G. and Croce,C.M.
  TITLE     Abnormalities at 14q32.1 in T cell malignancies involve two
            oncogenes
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2949-2951 (1999)
  MEDLINE   99178995
REFERENCE   2  (bases 1 to 1248)
  AUTHORS   Pekarsky,Y., Hallas,C. and Croce,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-1998) KCI, Thomas Jefferson University, 233 South
            10th Str. BLSB 1032, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .1248
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q12-q13"
     gene            163. .1248
                     /gene="TCL1B"
                     /note="T-cell leukemia/lymphoma 1B"
                     /pseudo
BASE COUNT      263 a    361 c    312 g    312 t
ORIGIN
```

Query Match 60.8%; Score 700.4; DB 9; Length 1248;
Best Local Similarity 81.5%; Pred. No. 1.2e-182;
Matches 928; Conservative 0; Mismatches 156; Indels 54; Gaps 8;

```
Qy      2 aggcgggtcccggttgcagacttgccatggcctccgaagcttctgtgcgtctaggggtgc 61
          ||||   || |  ||||| ||||||||||||||| ||||||||  || |||  |||  |
Db    149 AGGCGACCCCAGAGTGCAGCCTTGCCATGGCCTCCAAAGCTTCTCCGCATCTCAGGGCAC 208

Qy     62 cccctgg--ccgtctgtggatccagaggcctggcatctacgaagatgaggaggggagaac 119
          ||||     | ||||||||||  ||| ||||||||||| || |||||  ||||||| ||
Db    209 CCCCCCCAACTGTCTGTGGATCTGGAGACCTGGCATCTATGAGGATGACAAGGGGAGGAC 268

Qy    120 ctgggtgactgtggtcgtgcggttcaatccctcgcgtagggaatgggccagg------gc 173
          |||||||||||||| |||||  ||| ||||  ||  || ||| ||||||||
Db    269 CTGGGTGACTGTGGTTGTGCGCCTCAGTCCCCTGCACAGAGAAAGGGCCAGGTGCTCAAG 328

Qy    174 ctcccagggcagcagatatgaacccagcatcacagtgcacttgtggcagatggcagtgca 233
          |||||||||||||| ||||||||| |||||||||||||| ||||| ||||| |||||||
Db    329 CTCCCAGGGCAGCACATATGAACCTAGCATCACAGTGCACATGTGGTAGATGCCAGTGCA 388

Qy    234 tacccgggagctactctcctccggccagatgcccttctcccagctgcccgccgtgtggca 293
          ||||||||  | ||| ||| || ||||| |||||||||||||| ||||   ||||||||||
Db    389 CACCCGGGAACCACTGTCCCCCAGCCAGCTGCCCTTCTCCCACCTGCTTCCCGTGTGGCA 448

Qy    294 gctctacccccgggaggaagtaccgagcagcggattccagtttctgggaaatagcagacca 353
          ||||||||||  ||||| ||| | ||||||  ||||||||||||||||||||  ||
Db    449 GCTCTACCCCAAGAGGAGGTATCAAGCAGCAAATTCCAGTTTCTGGGAAATAGTGGA--- 505

Qy    354 tggccagattgactctatggagcagctggtcctaacatatcagccggagaggaaagactg 413
             |||||| | ||| ||||||||||||||||||||||||  |||||  |||||  ||||
Db    506 ---CCAGATCGTCTCCATGGAGCAGCTGGTCCTAACATATTGGCCGGCAAGGAATAACTG 562

Qy    414 acactgggagtggctggccctgctggccctgcctcttctggcctggtgtctcctcatgcc 473
          |                                ||| ||||||||  ||||||  ||  |||
Db    563 A------------------------------CTCCTCTGGCCTCATGTCTCTTCGGGCC 591
```

```
Qy    474 ccctcagtgaggatcttcatgtacctgctcttctgtttgcacacccagcatagcctcctt 533
          ||||||||||||||||  ||||| |||| ||| |||||||||||||| | ||||||||
Db    592 CCCTCAGTGAGGATCTTTGTGTACTTGCTATTCCGTTTGCACACCCAGCGTGGCCTCCTT 651

Qy    534 gcaggcagaaggcagtagggcccctgcacactcagtttctctcgttttccttagttatca 593
          |||||||| ||||||| |||||||||| ||||||| ||||||| ||||| | | ||||||
Db    652 GCAGGCAGGAGGCAGTGGGGCCCCTGCCCACTCAGCTTCTCTCATTTTCTTCACTTATCA 711

Qy    594 gtcctgtcctgtcccactcaggtctgtacttagggcagctggcctggatgggcttcact- 652
          ||| |||||||| |||||||| |||  ||| ||||||| ||||||||||||||||||||
Db    712 GTCTTGTCCTGTTCCACTCAGATCTACACTGAGGGCAGTTGGCCTGGATGGGCTTCACTA 771

Qy    653 ggggccctgtctgtgtgctgagccagtttcccctgctggctgcaagctgtgggttc--tt 710
          ||||||| |||||||  ||| ||| ||||||||||||||||||||||  |||||||  ||
Db    772 GGGGCCCCGTCTGTGCACTGGGCCCGTTTCCCCTGCTGGCTGCAAGCCATGGGTTCTTTT 831

Qy    711 tctcctctgtgcccctcatgctgatcttctagatgccactcccaaatcccctt-----ca 765
          |||||||| |||||||||||||| ||||||||||||||||||||||||||||     ||
Db    832 TCTCCTCTCTGCCCCTCATGCTGACCTTCTAGATGCCACTCCCAAATCCCCTTCACTCCA 891

Qy    766 tacccaccaggatgtgtgcccagccaggcctccagcacccccagtgcagctcgtgattgg 825
          |||||||||||  |   |||||| |||||||||  |||||| |||||||| | |||||||
Db    892 TACCCACCAGGCTTCATGCCCACCCAGGCCTCTGGCACCCTCAGTGCAGCCCATGATTGG 951

Qy    826 aaactcaccatcggcaggcagtggttcggtttaagagatggcattagagggagcccagtc 885
           ||||||||||| |||| |||||| |||||||||||||  |||   |||||||  |   |
Db    952 GAACTCACCATCAGCAGTCAGTGGCTCGGTTTAAGAGAGGGCCGCAGAGGGAACTGGGCT 1011

Qy    886 tggatgtggacttggatgccctgtgggtatcagttctgctgacactttggcccgaaatag 945
           |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1012 CGGATGTGGTCTTGGATGCCCTGTGGGTATCAGTTCTGCTGACACTTTGGCCTGAAATAG 1071

Qy    946 atccagtgctgagcaagcaatgtacaccggagcctcagtgagcccatctgcacagtgggg 1005
          ||  |||||||||||||| ||||  | |||||||||| || ||||||||||||||||||
Db   1072 ATAAAGTGCTGAGCAAGCAGTGTATGCTGGAGCCTCAGTAAGTCCATCTGCACAGTGGGG 1131

Qy   1006 agcatggagggatgggtttggcctgtgcttctgcttattcagtccttcagctcacggaag 1065
          ||| |||||||||||||||| | |||||||||| ||||||| |||||||||| | |
Db   1132 AGCGTGGAGGGATGGGTTTGGCTTATGCTTCTGCTTGTTCAGTCTTTCAGCTCACAGCA- 1190

Qy   1066 ggatgctagtccgtgaaggtgacctcacagtactggttaattaaacttttattgctcac 1123
          |||||||| ||| |||||||||||| |||||||||||||||| ||||||||||||
Db   1191 GGATGCTAATCCATGAAGGTGACCTCGCAGTACTGGTTAATTAAGCTTTATTGCTCAC 1248
```

```
RESULT  12
CNS01DX2
LOCUS       CNS01DX2   166308 bp    DNA             PRI       26-APR-2001
DEFINITION  Human chromosome 14 DNA sequence BAC R-164H13 of library RPCI-11
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION   AL139020
VERSION     AL139020.5  GI:13539193
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 166308)
  AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
  TITLE     Sequencing of the human chromosome 14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 166308)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On Apr 3, 2001 this sequence version replaced gi:12274871.
            ---------- Genome Center
            Center: Genoscope / Centre National de Sequencage
```

```
          Center code: GS
          Web site: http://www.genoscope.cns.fr/
          Contact: SeqRef@genoscope.cns.fr
          --------------
          The following BAC sequence is oriented from the T7 to the SP6 end.
          Upstream BAC (overlapping the T7 end) : R-1070N10
          Downstream BAC (overlapping the SP6 end) : R-185P18 (AC=AL133167)
          --------------- Summary Statistics
          Assembly program: Phrap; version 2.0
          Quality coverage: 6.03x in Q20 bases; sum-of-contigs
          ---------------
          ------------------------
          Overall quality chart :
          Range    :   bases
          0        :
          1  -  9  :       4
          10 - 19  :      53
          20 - 29  :     203
          30 - 39  :     622
          40 - 49  :    3717
          50 - 59  :    8215
          60 - 69  :   11927
          70 - 79  :   27324
          80 - 89  :   52859
          90 - 99  :   61384
          ------------------------
          Percentage of bases with a quality value >= 40 : 99 %.
FEATURES             Location/Qualifiers
     source          1. .166308
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="R-164H13"
                     /clone_lib="RPCI-11"
     STS             6803. .6993
                     /note="matching EMBL:R68409
                     RHdb:RH45107
                     dbSTS:STS38175
                     Identified using the e-PCR software (G. Schuler)"
     STS             9797. .10019
                     /note="matching EMBL:T96687
                     RHdb:RH45537
                     dbSTS:STS38605
                     Identified using the e-PCR software (G. Schuler)"
     STS             14657. .14855
                     /note="matching EMBL:R70888
                     RHdb:RH77821
                     dbSTS:STS54563
                     Identified using the e-PCR software (G. Schuler)"
     STS             16118. .16281
                     /note="matching EMBL:AA856757
                     RHdb:RH98896
                     dbSTS:STS68645
                     Identified using the e-PCR software (G. Schuler)"
     STS             46740. .46885
                     /note="matching EMBL:H01982
                     RHdb:RH44550
                     dbSTS:STS37619
                     Identified using the e-PCR software (G. Schuler)"
     STS             47277. .47485
                     /note="matching EMBL:X82240
                     RHdb:RH53532
                     dbSTS:STS26938
                     Identified using the e-PCR software (G. Schuler)"
BASE COUNT    46913 a  36249 c  37941 g  45205 t
ORIGIN


  Query Match             60.7%;  Score 699.6;  DB 9;  Length 166308;
  Best Local Similarity   99.4%;  Pred. No. 1.9e-182;
  Matches  702;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy    424 tggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcagtga 483
          |  || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28530 TCGCAGGCCCTGCTGGCCCTGCCTCTTCTGGCCTGGTGTCTCCTCATGCCCCCTCAGTGA 28589

Qy    484 ggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggcagaa 543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28590 GGATCTTCATGTACCTGCTCTTCTGTTTGCACACCCAGCATAGCCTCCTTGCAGGCAGAA 28649

Qy    544 ggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtcctgtcct 603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28650 GGCAGTAGGGCCCCTGCACACTCAGTTTCTCTCGTTTTCCTTAGTTATCAGTCCTGTCCT 28709

Qy    604 gtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggccctgtc 663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28710 GTCCCACTCAGGTCTGTACTTAGGGCAGCTGGCCTGGATGGGCTTCACTGGGGCCCTGTC 28769

Qy    664 tgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgtgcc 723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28770 TGTGTGCTGAGCCAGTTTCCCCTGCTGGCTGCAAGCTGTGGGTTCTTTCTCCTCTGTGCC 28829

Qy    724 cctcatgctgatcttctagatgccactcccaaatccccttcatacccaccaggatgtgtg 783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28830 CCTCATGCTGATCTTCTAGATGCCACTCCCAAATCCCCTTCATACCCACCAGGATGTGTG 28889

Qy    784 cccagccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggcagg 843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28890 CCCAGCCAGGCCTCCAGCACCCCCAGTGCAGCTCGTGATTGGAAACTCACCATCGGCAGG 28949

Qy    844 cagtggttcggtttaagagatggcattagagggagcccagtctggatgtggacttggatg 903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28950 CAGTGGTTCGGTTTAAGAGATGGCATTAGAGGGAGCCCAGTCTGGATGTGGACTTGGATG 29009

Qy    904 ccctgtgggtatcagttctgctgacactttggcccgaaatagatccagtgctgagcaagc 963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29010 CCCTGTGGGTATCAGTTCTGCTGACACTTTGGCCCGAAATAGATCCAGTGCTGAGCAAGC 29069

Qy    964 aatgtacaccggagcctcagtgagcccatctgcacagtggggagcatggagggatgggtt 1023
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29070 AATGTACACCAGAGCCTCAGTGAGCCCATCTGCACAGTGGGGAGCATGGAGGGATGGGTT 29129

Qy   1024 tggcctgtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtgaag 1083
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  29130 TGGCCTGTGCTTCTGCTTATTCAGTCCTTCAGCTCACGGAAGGGATGCTAGTCCGTGAAG 29189

Qy   1084 gtgacctcacagtactggttaattaaactttattgctcactgtcaa 1129
          ||||||||||||||||||||||||||||||||||||||||||| |
Db  29190 GTGACCTCACAGTACTGGTTAATTAAACTTTATTGCTCACTGTCCA 29235


RESULT  13
CNS01DV7
LOCUS       CNS01DV7    210791 bp    DNA             PRI       23-MAY-2001
DEFINITION  Human chromosome 14 DNA sequence BAC R-1070N10 of library RPCI-11
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION   AL133467
VERSION     AL133467.4  GI:14268349
KEYWORDS    HTG; HTGS_ACTIVEFIN.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 210791)
  AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
  TITLE     Sequencing of the human chromosome 14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 210791)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAY-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
```

```
COMMENT     On May 30, 2001 this sequence version replaced gi:12657316.
            -------------- Genome Center
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: SeqRef@genoscope.cns.fr
            ---------------
            The following BAC sequence is oriented from the T7 to the SP6 end.
            Upstream BAC (overlapping the T7 end) : C-2588C21
            Downstream BAC (overlapping the SP6 end) : R-164H13 (AC=AL139020)
            -------------- Summary Statistics
            Assembly program: Phrap; version 2.0
            Quality coverage: 7.83x in Q20 bases; sum-of-contigs
            ---------------
            ------------------------
            Overall quality chart :
            Range    :   bases
            0        :
            1  -  9  :
            10 - 19  :
            20 - 29  :       3
            30 - 39  :      34
            40 - 49  :    1491
            50 - 59  :    5034
            60 - 69  :    7834
            70 - 79  :   22859
            80 - 89  :   63567
            90 - 99  :  109969
            ------------------------
            Percentage of bases with a quality value >= 40 : 99 %.
FEATURES              Location/Qualifiers
     source           1. .210791
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /chromosome="14"
                      /clone="R-1070N10"
                      /clone_lib="RPCI-11"
     STS              11396. .11567
                      /note="matching EMBL:H04638
                      RHdb:RH44553
                      dbSTS:STS37622
                      Identified using the e-PCR software (G. Schuler)"
     STS              11476. .11602
                      /note="matching EMBL:R21904
                      RHdb:RH69831
                      dbSTS:STS49708
                      Identified using the e-PCR software (G. Schuler)"
     STS              34400. .34562
                      /note="matching EMBL:Z16446
                      RHdb:RH95830
                      RHdb:RH48997
                      RHdb:RH12943
                      RHdb:RH785
                      dbSTS:STS23702
                      Identified using the e-PCR software (G. Schuler)"
     STS              44291. .44412
                      /note="matching EMBL:AA857911
                      RHdb:RH93740
                      dbSTS:STS66733
                      Identified using the e-PCR software (G. Schuler)"
     STS              55616. .55844
                      /note="matching EMBL:T72567
                      RHdb:RH53712
                      dbSTS:STS42889
                      Identified using the e-PCR software (G. Schuler)"
     STS              55639. .55818
                      /note="matching EMBL:T72567
                      RHdb:RH44140
                      RHdb:RH1627
                      dbSTS:STS4726
                      Identified using the e-PCR software (G. Schuler)"
     STS              156248. .156397
                      /note="matching EMBL:AA129468
                      RHdb:RH48023
                      dbSTS:STS41082
                      Identified using the e-PCR software (G. Schuler)"
     STS              181393. .181583
                      /note="matching EMBL:R68409
                      RHdb:RH45107
                      dbSTS:STS38175
                      Identified using the e-PCR software (G. Schuler)"
     STS              184387. .184609
                      /note="matching EMBL:T96687
                      RHdb:RH45537
                      dbSTS:STS38605
                      Identified using the e-PCR software (G. Schuler)"
     STS              189248. .189446
                      /note="matching EMBL:R70888
                      RHdb:RH77821
                      dbSTS:STS54563
                      Identified using the e-PCR software (G. Schuler)"
     STS              190708. .190871
                      /note="matching EMBL:AA856757
                      RHdb:RH98896
                      dbSTS:STS68645
                      Identified using the e-PCR software (G. Schuler)"
BASE COUNT    55743 a  48482 c  50054 g  56512 t
ORIGIN


  Query Match           60.7%;  Score 699.6;  DB 9;  Length 210791;
  Best Local Similarity 99.4%;  Pred. No. 1.9e-182;
  Matches  702;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     424 tggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcagtga 483
           | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203113 TCGCAGGCCCTGCTGGCCCTGCCTCTTCTGGCCTGGTGTCTCCTCATGCCCCCTCAGTGA 203172

Qy     484 ggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggcagaa 543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203173 GGATCTTCATGTACCTGCTCTTCTGTTTGCACACCCAGCATAGCCTCCTTGCAGGCAGAA 203232

Qy     544 ggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtcctgtcct 603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203233 GGCAGTAGGGCCCCTGCACACTCAGTTTCTCTCGTTTTCCTTAGTTATCAGTCCTGTCCT 203292

Qy     604 gtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggccctgtc 663
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203293 GTCCCACTCAGGTCTGTACTTAGGGCAGCTGGCCTGGATGGGCTTCACTGGGGCCCTGTC 203352

Qy     664 tgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgtgcc 723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203353 TGTGTGCTGAGCCAGTTTCCCCTGCTGGCTGCAAGCTGTGGGTTCTTTCTCCTCTGTGCC 203412

Qy     724 cctcatgctgatcttctagatgccactcccaaatccccttcatacccaccaggatgtgtg 783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203413 CCTCATGCTGATCTTCTAGATGCCACTCCCAAATCCCCTTCATACCCACCAGGATGTGTG 203472

Qy     784 cccagccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggcagg 843
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203473 CCCAGCCAGGCCTCCAGCACCCCCAGTGCAGCTCGTGATTGGAAACTCACCATCGGCAGG 203532

Qy     844 cagtggttcggtttaagagatggcattagagggagcccagtctggatgtggacttggatg 903
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203533 CAGTGGTTCGGTTTAAGAGATGGCATTAGAGGGAGCCCAGTCTGGATGTGGACTTGGATG 203592

Qy     904 ccctgtgggtatcagttctgctgacactttggcccgaaatagatccagtgctgagcaagc 963
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203593 CCCTGTGGGTATCAGTTCTGCTGACACTTTGGCCCGAAATAGATCCAGTGCTGAGCAAGC 203652

Qy     964 aatgtacaccggagcctcagtgagcccatctgcacagtggggagcatggagggatgggtt 1023
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  203653 AATGTACACCAGAGCCTCAGTGAGCCCATCTGCACAGTGGGGAGCATGGAGGGATGGGTT 203712

Qy    1024 tggcctgtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtgaag 1083
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 203713 TGGCCTGTGCTTCTGCTTATTCAGTCCTTCAGCTCACGGAAGGGATGCTAGTCCGTGAAG 203772

Qy   1084 gtgacctcacagtactggttaattaaacttttattgctcactgtcaa 1129
          |||||||||||||||||||||||||||||||||||||||||||||| |
Db 203773 GTGACCTCACAGTACTGGTTAATTAAACTTTATTGCTCACTGTCCA 203818


RESULT  14
AC016603
LOCUS       AC016603   110879 bp    DNA             PRI       20-JUL-2001
DEFINITION  Homo sapiens chromosome 5 clone CTD-2132G23, complete sequence.
ACCESSION   AC016603
VERSION     AC016603.6  GI:14971185
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110879)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 110879)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 110879)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 20, 2001 this sequence version replaced gi:7711572.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
            STS Content:
            WI-31186 G24491.
FEATURES             Location/Qualifiers
     source          1. .110879
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2132G23"
BASE COUNT    33303 a  18836 c  20919 g  37821 t
ORIGIN


  Query Match             60.4%;  Score 695.4;  DB 9;  Length 110879;
  Best Local Similarity   82.0%;  Pred. No. 2.7e-181;
  Matches  905;  Conservative    0;  Mismatches  146;  Indels   52;  Gaps    7;

Qy      51 tctaggggtgccccctggccgtctgtggatccagaggcctggcatctacgaagatgagga 110
           ||  ||    |||||   | ||||||||||  ||| ||||||||||| || |||||  |
Db      16 TCAGGGCACCCCCCCCCAACTGTCTGTGGATCTGGAGACCTGGCATCTATGAGGATGACAA 75

Qy     111 ggggagaacctgggtgactgtggtcgtgcggttcaatccctcgcgtagggaatgggccag 170
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76 GGGGAGGACCTGGGTGACTGTGGTTGTGCGCCTCAGTCCCCTGCACAGAGAAAGGGCCAG 135

Qy     171 g------gcctcccagggcagcagatatgaacccagcatcacagtgcacttgtggcagat 224
           |      ||||||||||||||| |||||||||| ||||||||||||||| ||||| ||||
Db     136 GTGCTCAAGCTCCCAGGGCAGCACATATGAACCTAGCATCACAGTGCACATGTGGTAGAT 195

Qy     225 ggcagtgcatacccgggagctactctcctccggccagatgcccttctcccagctgcccgc 284
           | ||||||| |||||||| | ||| ||| || ||||| |||||||||||||| ||||   |
Db     196 GCCAGTGCACACCCGGGAACCACTGTCCCCCAGCCAGCTGCCCTTCTCCCACCTGCTTCC 255

Qy     285 cgtgtggcagctctaccccgggaggaagtaccgagcagcggattccagtttctgggaaat 344
           ||||||||||||||||||  ||||| ||| | ||||||  ||||||||||||||||||||

Db     256 CGTGTGGCAGCTCTACCCCAAGAGGAGGTATCAAGCAGCAAATTCCAGTTTCTGGGAAAT 315

Qy     345 agcagaccatggccagattgactctatggagcagctggtcctaacatatcagccggagag 404
           ||  |||||      ||| | ||| ||||||||||||||||||||||  |||||  ||
Db     316 AGTGGACCA------GATCGTCTCCATGGAGCAGCTGGTCCTAACATATTGGCCGGCAAG 369

Qy     405 gaaagactgacactgggagtggctggccctgctggccctgcctcttctggcctggtgtct 464
           |||  |||||                                  ||| ||||||||  |||||
Db     370 GAATAACTGA------------------------------CTCCTCTGGCCTCATGTCT 398

Qy     465 cctcatgccccctcagtgaggatcttcatgtacctgctcttctgtttgcacacccagcat 524
           | ||  ||||||||||||||||||||  ||||| |||| ||| ||||||||||||||| |
Db     399 CTTCGGGCCCCCTCAGTGAGGATCTTTGTGTACTTGCTATTCCGTTTGCACACCCAGCGT 458

Qy     525 agcctccttgcaggcagaaggcagtagggcccctgcacactcagtttctctcgttttcct 584
           ||||||||||||||||| |||||||| |||||||||| |||||| ||||||| ||||| |
Db     459 GGCCTCCTTGCAGGCAGGAGGCAGTGGGGCCCCTGCCCACTCAGCTTCTCTCATTTTCTT 518

Qy     585 tagttatcagtcctgtcctgtcccactcaggtctgtacttagggcagctggcctggatgg 644
           | |||||||| |||||||| |||||||| |||  ||| ||||||| ||||||||||||
Db     519 CACTTATCAGTCTTGTCCTGTTCCACTCAGATCTACACTGAGGGCAGTTGGCCTGGATGG 578

Qy     645 gcttcact-ggggccctgtctgtgtgctgagccagtttcccctgctggctgcaagctgtg 703
           |||||||| ||||||| |||||||  ||| ||| |||||||||||||||||||||||  ||
Db     579 GCTTCACTAGGGGCCCCGTCTGTGCACTGGGCCCGTTTCCCCTGCTGGCTGCAAGCCATG 638

Qy     704 ggttc--tttctcctctgtgcccctcatgctgatcttctagatgccactcccaaatcccc 761
           |||||  |||||||||| |||||||||||||| |||||||||||||||||||||||||||
Db     639 GGTTCTTTTTCTCCTCTCTGCCCCTCATGCTGACCTTCTAGATGCCACTCCCAAATCCCC 698

Qy     762 tt-----catacccaccaggatgtgtgcccagccaggcctccagcacccccagtgcagct 816
           ||     |||||||||||| |   |||||| |||||||||  |||||| |||||||||
Db     699 TTCACTCCATACCCACCAGGCTTCATGCCCACCCAGGCCTCTGGCACCCTCAGTGCAGCC 758

Qy     817 cgtgattggaaactcaccatcggcaggcagtggttcggtttaagagatggcattagaggg 876
           | ||||||| |||||||||| |||| |||||| |||||||||||| |||   ||||||
Db     759 CATGATTGGGAACTCACCATCAGCAGTCAGTGGCTCGGTTTAAGAGAGGGCCGCAGAGGG 818

Qy     877 agcccagtctggatgtggacttggatgccctgtgggtatcagttctgctgacactttggc 936
           | |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     819 AACTGGGTCTGGATGTGGACTTGGATGCCCTGTGGGTATCAGTTCTGCTGACACTTTGGC 878

Qy     937 ccgaaatagatccagtgctgagcaagcaatgtacaccggagcctcagtgagcccatctgc 996
           | |||||||||  |||||||||||||||| ||||  | |||||||||| || ||||||||
Db     879 CTGAAATAGATAAAGTGCTGAGCAAGCAGTGTATGCTGGAGCCTCAGTAAGTCCATCTGC 938

Qy     997 acagtggggagcatggagggatgggtttggcctgtgcttctgcttattcagtccttcagc 1056
           |||||||||| |||||||||||||||||| | |||||||||| ||||||| ||||||
Db     939 ACAGTGGGGAGCGTGGAGGGATGGGTTTGGCTTATGCTTCTGCTTGTTCAGTCTTTCAGC 998

Qy    1057 tcacggaagggatgctagtccgtgaaggtgacctcacagtactggttaattaaacttttat 1116
           |||| | | |||||||| ||| |||||||||||||| |||||||||||||||| ||||||
Db     999 TCACAGCA-GGATGCTAATCCATGAAGGTGACCTCGCAGTACTGGTTAATTAAGCTTTAT 1057

Qy    1117 tgctcactgtcaaaaaaaaaaaa 1139
           |||||||||  ||| ||||||||
Db    1058 TGCTCACTGAAAAAGAAAAAAAA 1080


RESULT  15
AB025272S3
LOCUS       AB025272S3   1717 bp    DNA             PRI       14-APR-2000
DEFINITION  Homo sapiens TML1 gene for TCL1 / MTCP1-like 1, exon 3 and complete
            cds.
ACCESSION   AB025274
VERSION     AB025274.1  GI:5478618
KEYWORDS    TML1; TCL1 / MTCP1-like 1.
SEGMENT     3 of 3
SOURCE      Homo sapiens DNA, clone:pCOS231.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
```

```
  AUTHORS   Sugimoto,J., Hatakeyama,T., Narducci,M.G., Russo,G. and Isobe,M.
  TITLE     Identification of the TCL1/MTCP1-like 1 (TML1) gene from the region
            next to the TCL1 locus
  JOURNAL   Cancer Res. 59 (10), 2313-2317 (1999)
  MEDLINE   99274532
REFERENCE   2  (bases 1 to 1717)
  AUTHORS   Isobe,M., Sugimoto,J. and Hatakeyama,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-1999) to the DDBJ/EMBL/GenBank databases.
            Masaharu Isobe, Toyama University, Faculty of Engineering; Gofuku
            3160, Toyama, Toyama 930-8555, Japan
            (E-mail:isobe@eng.toyama-u.ac.jp, Tel:+81-76-445-6872,
            Fax:+81-76-445-6874)
FEATURES             Location/Qualifiers
     source          1. .1717
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="pCOS231"
                     /map="14q32.1"
     gene            join(AB025272.1:119. .310,AB025273.1:223. .393,265. .318)
                     /gene="TML1"
     CDS             join(AB025272.1:149. .310,AB025273.1:223. .393,265. .318)
                     /gene="TML1"
                     /codon_start=1
                     /product="TCL1 / MTCP1-like 1"
                     /protein_id="BAA82476.1"
                     /db_xref="GI:5478620"
                     /translation="MASEASVRLGVPPGRLWIQRPGIYEDEEGRTWVTVVVRFNPSRR
                     EWARASQGSRYEPSITVHLWQMAVHTRELLSSGQMPFSQLPAVWQLYPRRKYRAADSS
                     FWEIADHGQIDSMEQLVLTYQPERKD"
     exon            265. .1623
                     /number=3
                     /product="TCL1 / MTCP1-like 1 (TML1)"
     polyA_signal    1601. .1606
BASE COUNT      340 a    466 c    473 g    438 t
ORIGIN


  Query Match           59.7%;  Score 687.6;  DB 9;  Length 1717;
  Best Local Similarity 99.3%;  Pred. No. 4.3e-179;
  Matches  701;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy     424 tggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcagtga 483
           | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     921 TCGCAGGCCCTGCTGGCCCTGCCTCTTCTGGCCTGGTGTCTCCTCATGCCCCCTCAGTGA 980

Qy     484 ggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggcagaa 543
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     981 GGATCTTCATGTACCTGCTCTTCTGTTTGCACACCCAGCATAGCCT-CTTGCAGGCAGAA 1039

Qy     544 ggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtcctgtcct 603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1040 GGCAGTAGGGCCCCTGCACACTCAGTTTCTCTCGTTTTCCTTAGTTATCAGTCCTGTCCT 1099

Qy     604 gtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggccctgtc 663
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1100 GTCCCACTCAGGTCTGTACTTAGGGCAGCTGGCCTGGATGGGCTTCACTGGGGCCCTGTC 1159

Qy     664 tgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgtgcc 723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1160 TGTGTGCTGAGCCAGTTTCCCCTGCTGGCTGCAAGCTGTGGGTTCTTTCTCCTCTGTGCC 1219

Qy     724 cctcatgctgatcttctagatgccactcccaaatccccttcatacccaccaggatgtgtg 783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1220 CCTCATGCTGATCTTCTAGATGCCACTCCCAAATCCCCTTCATACCCACCAGGATGTGTG 1279

Qy     784 cccagccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggcagg 843
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1280 CCCAGCCAGGCCTCCAGCACCCCCAGTGCAGCTCGTGATTGGAAACTCACCATCGGCAGG 1339

Qy     844 cagtggttcggtttaagagatggcattagagggagcccagtctggatgtggacttggatg 903
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    1340 CAGTGGTTCGGTTTAAGAGATGGCATTAGAGGGAGCCCAGTCTGGATGTGGACTTGGATG 1399

Qy     904 ccctgtgggtatcagttctgctgacactttggcccgaaatagatccagtgctgagcaagc 963
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1400 CCCTGTGGGTATCAGTTCTGCTGACACTTTGGCCCGAAATAGATCCAGTGCTGAGCAAGC 1459

Qy     964 aatgtacaccggagcctcagtgagcccatctgcacagtggggagcatggagggatgggtt 1023
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1460 AATGTACACCAGAGCCTCAGTGAGCCCATCTGCACAGTGGGGAGCATGGAGGGATGGGTT 1519

Qy    1024 tggcctgtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtgaag 1083
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1520 TGGCCTGTGCTTCTGCTTATTCAGTCCTTCAGCTCACGGAAGGGATGCTAGTCCGTGAAG 1579

Qy    1084 gtgacctcacagtactggttaattaaactttattgctcactgtcaa 1129
           |||||||||||||||||||||||||||||||||||||||||||| |
Db    1580 GTGACCTCACAGTACTGGTTAATTAAACTTTATTGCTCACTGTCCA 1625


Search completed: November 29, 2001, 08:52:45
Job time: 17206 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 04:06:44 ; Search time 576.65 Seconds
(without alignments)
1712.718 Million cell updates/sec

Title: US-09-526-329-38
Perfect score: 1152
Sequence: 1 gaggcgggtcccggttgcag.....aaaaaaaaaaaaaaaaaaa 1152

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1152	100.0	1152	21	AAA75822	Nucleotide sequenc
2	701.2	60.9	6486	21	AAA75823	Nucleotide sequenc
3	393.8	34.2	697	22	AAH98245	Human EST-derived
4	62.8	5.5	561	21	AAA89670	Human T cell leuka
5	62.8	5.5	1235	22	AAI60765	Human polynucleoti
6	62.8	5.5	1324	17	AAT18876	Human TCL-1 cDNA.
7	62.8	5.5	1368	22	AAI58979	Human polynucleoti
8	50	4.3	4922	17	AAT18877	Human TCL-1 gene.
c 9	45.6	4.0	10732	21	AAA10594	Gene encoding a su
10	41.6	3.6	1533	20	AAV84490	Human secreted pro
11	41.6	3.6	1719	22	AAD04477	Human sphingosine
12	41	3.6	1512	21	AAC75966	Human ORFX ORF1521
13	40.4	3.5	2440	21	AAC77441	Human ORFX ORF2996
14	39.8	3.5	2617	22	AAH34812	Human colon cancer
15	39.2	3.4	693	21	AAA77669	Human PRO1286 cDNA
16	39.2	3.4	693	21	AAA37065	Human PRO1286 (UNQ
17	39.2	3.4	693	22	AAS21493	Human cDNA sequenc
18	39.2	3.4	693	22	AAF54301	DNA encoding prote
19	39.2	3.4	693	22	AAC91565	Human PRO1286 cDNA
20	38.6	3.4	284	20	AAZ24580	Human lung tumor a
21	38.6	3.4	284	21	AAC65819	Human lung cancer-
22	38.6	3.4	496	21	AAC78292	Human cancer assoc
23	38.4	3.3	2468	22	AAH33714	Human colon cancer
24	38.2	3.3	798	20	AAX84203	DNA encoding human
25	38.2	3.3	798	21	AAC79432	cDNA sequence of h
26	38	3.3	1097	22	AAF61108	P. patens S-adenos
27	38	3.3	1896	21	AAC74297	Human secreted pro
28	37.8	3.3	585	20	AAX37386	Human secreted pro
29	37.8	3.3	887	19	AAV41915	Nucleotide sequenc
30	37.8	3.3	1314	21	AAC60057	Human secreted pro
31	37.8	3.3	1778	22	AAF54816	Nucleotide sequenc
32	37.8	3.3	1797	20	AAX85038	Human secreted pro
33	37.8	3.3	4274	15	AAQ79407	Human NMDAR2 recep
34	37.8	3.3	4289	15	AAQ79405	Human NMDAR2 recep
35	37.8	3.3	4325	15	AAQ79404	Human NMDAR2 recep
36	37.8	3.3	4340	15	AAQ79372	Human N-methyl-D-a
37	37.8	3.3	4349	15	AAQ79406	Human NMDAR2 recep
38	37.8	3.3	4364	15	AAQ79403	Human NMDAR2 recep
39	37.6	3.3	395	22	AAH33388	Human colon cancer
40	37.6	3.3	998	11	AAQ04796	Sequence encoding
41	37.6	3.3	1989	21	AAA51408	A. thaliana LDV115
42	37.6	3.3	2017	21	AAA97355	Human colorectal c
43	37.6	3.3	4064	21	AAA07587	Mouse piwi gene, d
44	37.4	3.2	325	21	AAC99965	Human secreted pro
45	37.4	3.2	550	21	AAC93590	Cat flea head and

ALIGNMENTS

RESULT 1
AAA75822
ID AAA75822 standard; DNA; 1152 BP.
XX
AC AAA75822;
XX
DT 22-JAN-2001 (first entry)
XX
DE Nucleotide sequence of the human Tcl-1b cDNA.
XX
KW Tcl-1; Tcl-1b; T cell malignancy; chromosome 14 abnormality; lymphoma;
KW T-cell leukaemia; immunodeficiency syndrome; ataxia-teleangiectasia; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 28..414
FT /*tag= a
FT /product= "Tcl-1b protein"
XX
PN WO200055169-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06612.
XX
PR 15-MAR-1999; 99US-0124714.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Croce CM, Pekarsky Y;
XX
DR WPI; 2000-611514/58.
DR P-PSDB; AAB18763.

```
XX
PT   Novel nucleic acid of Tcl-1 gene family, Tcl-1b, expressed in low
PT   levels in normal bone marrow and peripheral lymphocytes, but activated
PT   in T-cell leukaemia and lymphoma, used to identify chromosome 14
PT   abnormalities  -
XX
PS   Claim 5; Page 63; 70pp; English.
XX
CC   The present sequence encodes a human Tcl-1b protein. The Tcl-1b gene
CC   is implicated in the development of T cell malignancies. Fragments of
CC   Tcl-1b cDNA sequences are used for detecting a target sequence
CC   indicating a chromosome 14 abnormality, such as a (14:14)(q11:q32)
CC   translocation or a (14)(q11:q32) inversion. Tcl-1b antisense sequences
CC   and antibodies are useful for treating a disease state such as T-cell
CC   leukaemia or lymphoma associated with a chromosome 14 abnormality. The
CC   Tcl-1b gene and its gene product are useful for treating disease states
CC   associated with the Tcl-1b locus on chromosome protein including
CC   T-prolymphocytic leukaemias, acute and chronic leukaemias associated
CC   with the immunodeficiency syndrome ataxia-teleangiectasia (AT).
XX
SQ   Sequence 1152 BP; 232 A; 325 C; 324 G; 271 T; 0 other;


  Query Match             100.0%;  Score 1152;  DB 21;  Length 1152;
  Best Local Similarity   100.0%;  Pred. No. 9.2e-299;
  Matches 1152;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gaggcgggtcccggttgcagacttgccatggcctccgaagcttctgtgcgtctaggggtg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gaggcgggtcccggttgcagacttgccatggcctccgaagcttctgtgcgtctaggggtg 60

Qy     61 ccccctggccgtctgtggatccagaggcctggcatctacgaagatgaggaggggagaacc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ccccctggccgtctgtggatccagaggcctggcatctacgaagatgaggaggggagaacc 120

Qy    121 tgggtgactgtggtcgtgcggttcaatccctcgcgtagggaatgggccagggcctcccag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 tgggtgactgtggtcgtgcggttcaatccctcgcgtagggaatgggccagggcctcccag 180

Qy    181 ggcagcagatatgaacccagcatcacagtgcacttgtggcagatggcagtgcatacccgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ggcagcagatatgaacccagcatcacagtgcacttgtggcagatggcagtgcatacccgg 240

Qy    241 gagctactctcctccggccagatgcccttctcccagctgcccgccgtgtggcagctctac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gagctactctcctccggccagatgcccttctcccagctgcccgccgtgtggcagctctac 300

Qy    301 cccgggaggaagtaccgagcagcggattccagtttctgggaaatagcagaccatggccag 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 cccgggaggaagtaccgagcagcggattccagtttctgggaaatagcagaccatggccag 360

Qy    361 attgactctatggagcagctggtcctaacatatcagccggagaggaaagactgacactgg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 attgactctatggagcagctggtcctaacatatcagccggagaggaaagactgacactgg 420

Qy    421 gagtggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcag 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gagtggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcag 480

Qy    481 tgaggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 tgaggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggca 540

Qy    541 gaaggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtcctgt 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gaaggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtcctgt 600

Qy    601 cctgtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggccct 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 cctgtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggccct 660

Qy    661 gtctgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgt 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gtctgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgt 720

Qy    721 gcccctcatgctgatcttctagatgccactcccaaatcccctttcatacccaccaggatgt 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gcccctcatgctgatcttctagatgccactcccaaatcccctttcatacccaccaggatgt 780

Qy    781 gtgcccagccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggc 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gtgcccagccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggc 840

Qy    841 aggcagtggttcggtttaagagatggcattagagggagcccagtctggatgtggacttgg 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 aggcagtggttcggtttaagagatggcattagagggagcccagtctggatgtggacttgg 900

Qy    901 atgccctgtgggtatcagttctgctgacactttggcccgaaatagatccagtgctgagca 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 atgccctgtgggtatcagttctgctgacactttggcccgaaatagatccagtgctgagca 960

Qy    961 agcaatgtacaccggagcctcagtgagcccatctgcacagtggggagcatggagggatgg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 agcaatgtacaccggagcctcagtgagcccatctgcacagtggggagcatggagggatgg 1020

Qy   1021 gtttggcctgtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gtttggcctgtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtg 1080

Qy   1081 aaggtgacctcacagtactggttaattaaactttattgctcactgtcaaaaaaaaaaaaa 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 aaggtgacctcacagtactggttaattaaactttattgctcactgtcaaaaaaaaaaaaa 1140

Qy   1141 aaaaaaaaaaaa 1152
          ||||||||||||
Db   1141 aaaaaaaaaaaa 1152


RESULT   2
AAA75823
ID   AAA75823 standard; DNA; 6486 BP.
XX
AC   AAA75823;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Nucleotide sequence of the human Tcl-1b gene.
XX
KW   Tcl-1; Tcl-1b; T cell malignancy; chromosome 14 abnormality; lymphoma;
KW   T-cell leukaemia; immunodeficiency syndrome; ataxia-teleangiectasia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055169-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06612.
XX
PR   15-MAR-1999;   99US-0124714.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Croce CM,  Pekarsky Y;
XX
DR   WPI; 2000-611514/58.
XX
PT   Novel nucleic acid of Tcl-1 gene family, Tcl-1b, expressed in low
PT   levels in normal bone marrow and peripheral lymphocytes, but activated
PT   in T-cell leukaemia and lymphoma, used to identify chromosome 14
PT   abnormalities  -
XX
PS   Claim 12; Page 64-67; 70pp; English.
XX
CC   The present sequence encodes a human Tcl-1b protein. The Tcl-1b gene
CC   is implicated in the development of T cell malignancies. Fragments of
```

```
CC   Tcl-1b cDNA sequences are used for detecting a target sequence
CC   indicating a chromosome 14 abnormality, such as a (14:14)(q11:q32)
CC   translocation or a (14)(q11:q32) inversion. Tcl-1b antisense sequences
CC   and antibodies are useful for treating a disease state such as T-cell
CC   leukaemia or lymphoma associated with a chromosome 14 abnormality. The
CC   Tcl-1b gene and its gene product are useful for treating disease states
CC   associated with the Tcl-1b locus on chromosome protein including
CC   T-prolymphocytic leukaemias, acute and chronic leukaemias associated
CC   with the immunodeficiency syndrome ataxia-teleangiectasia (AT).
XX
SQ   Sequence 6486 BP; 1600 A; 1559 C; 1687 G; 1639 T; 1 other;

  Query Match             60.9%;  Score 701.2;  DB 21;  Length 6486;
  Best Local Similarity   99.6%;  Pred. No. 1e-177;
  Matches  703;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     424 tggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcagtga 483
           | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5628 tcgcaggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcagtga 5687

Qy     484 ggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggcagaa 543
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5688 ggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggcagaa 5747

Qy     544 ggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtcctgtcct 603
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5748 ggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtcctgtcct 5807

Qy     604 gtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggccctgtc 663
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5808 gtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggccctgtc 5867

Qy     664 tgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgtgcc 723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5868 tgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcctctgtgcc 5927

Qy     724 cctcatgctgatcttctagatgccactcccaaatccccttcatacccaccaggatgtgtg 783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5928 cctcatgctgatcttctagatgccactcccaaatccccttcatacccaccaggatgtgtg 5987

Qy     784 cccagccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggcagg 843
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5988 cccagccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggcagg 6047

Qy     844 cagtggttcggtttaagagatggcattagagggagcccagtctggatgtggacttggatg 903
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6048 cagtggttcggtttaagagatggcattagagggagcccagtctggatgtggacttggatg 6107

Qy     904 ccctgtgggtatcagttctgctgacactttggcccgaaatagatccagtgctgagcaagc 963
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6108 ccctgtgggtatcagttctgctgacactttggcccgaaatagatccagtgctgagcaagc 6167

Qy     964 aatgtacaccggagcctcagtgagcccatctgcacagtggggagcatggagggatgggtt 1023
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6168 aatgtacaccggagcctcagtgagcccatctgcacagtggggagcatggagggatgggtt 6227

Qy    1024 tggcctgtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtgaag 1083
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6228 tggcctgtgcttctgcttattcagtccttcagctcacggaagggatgctagtccgtgaag 6287

Qy    1084 gtgacctcacagtactggttaattaaaactttattgctcactgtcaa 1129
           |||||||||||||||||||||||||||||||||||||||||||| |
Db    6288 gtgacctcacagtactggttaattaaaactttattgctcactgtcca 6333


RESULT   3
AAH98245
ID   AAH98245 standard; cDNA; 697 BP.
XX
AC   AAH98245;
XX
DT   12-OCT-2001  (first entry)

XX
DE   Human EST-derived coding sequence SEQ ID NO: 102.
XX
KW   Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW   tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW   diagnostics; forensic test; gene mapping; genetic disorder;
KW   biodiversity; gene therapy; nutrition; ss.
XX
OS   Homo sapiens.
XX
PN   WO200154477-A2.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-US02687.
XX
PR   25-JAN-2000; 2000US-0491404.
PR   17-JUL-2000; 2000US-0617746.
PR   03-AUG-2000; 2000US-0631451.
PR   15-SEP-2000; 2000US-0663870.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Zhou P,  Qian XB,  Wang Z,  Chen R,  Asundi V;
PI   Cao Y,  Drmanac RA,  Zhang J,  Werhman T;
XX
DR   WPI; 2001-476164/51.
DR   P-PSDB; AAM23586.
XX
PT   Isolated polypeptide for treatment of diseases, diagnostics, raising
PT   antibodies and research use -
XX
PS   Claim 1; Page 246; 1275pp; English.
XX
CC   The present invention provides the protein and coding sequences of novel
CC   proteins from a variety of organisms, including human, dog, cat, horse,
CC   cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC   urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC   from the organism of interest. They can be used in diagnostics,
CC   forensics, gene mapping, identification of mutations, to assess
CC   biodiversity and for nutritional purposes. The present sequence is a cDNA
CC   of the invention.
XX
SQ   Sequence 697 BP; 135 A; 204 C; 181 G; 175 T; 2 other;

  Query Match             34.2%;  Score 393.8;  DB 22;  Length 697;
  Best Local Similarity   79.6%;  Pred. No. 8.4e-96;
  Matches  569;  Conservative    0;  Mismatches   99;  Indels   47;  Gaps    7;

Qy     316 cgagcagcggattccagtttctgggaaatagcagaccatggccagattgactctatggag 375
           | |||||| ||||||||||||||||||||||  ||      |||||| | ||| ||||||
Db       1 caagcagcaaattccagtttctgggaaatagtgga------ccagatcgtctccatggag 54

Qy     376 cagctggtcctaacatatcagccggagaggaaagactgacactgggagtggctggccctg 435
           |||||||||||||||||| ||||| |||||  |||||
Db      55 cagctggtcctaacatattggccggcaaggaataactga---------------------- 93

Qy     436 ctggccctgcctcttctggcctggtgtctcctcatgccccctcagtgaggatcttcatgt 495
                   ||| ||||||||  ||||||  ||  ||||||||||||||||||||  |||
Db      94 ----------ctcctctggcctcatgtctcttcgggccccctcagtgaggatctttgtgt 143

Qy     496 acctgctcttctgtttgcacacccagcatagcctccttgcaggcagaaggcagtagggcc 555
           || |||| ||| |||||||||||||| | ||||||||||||||| ||||||| |||||
Db     144 acttgctattccgtttgcacacccagcgtggcctccttgcaggcaggaggcagtggggcc 203

Qy     556 cctgcacactcagtttctctcgttttccttagttatcagtcctgtcctgtcccactcagg 615
           ||||| ||||||| ||||||| |||| | | |||||||| |||||||| |||||||
Db     204 cctgcccactcagcttctctcattttcttcacttatcagtcttgtcctgttccactcaaa 263

Qy     616 tctgtacttagggcagctggcctggatgggcttcact-ggggccctgtctgtgtgctgag 674
           |||  ||| ||||||| |||||||||||||||||| ||||||| |||||||  ||| |
Db     264 tctacactgagggcagttggcctggatgggcttcactaggggccccgtctgtgcactggg 323
```

```
Qy    675 ccagtttcccctgctggctgcaagctgtgggttc--tttctcctctgtgcccctcatgct 732
          || ||||||||||||||||||||||  |||||||  ||||||||||  ||||||||||||
Db    324 cccgtttcccctgctggctgcaagccatgggttctttttctcctctctgcccctcatgct 383

Qy    733 gatcttctagatgccactcccaaatcccctt-----catacccaccaggatgtgtgccca 787
          || |||||||||||||||||||||||||||     ||||||||||||  |   ||||||
Db    384 gaccttctagatgccactcccaaatcccctttcactccatacccaccaggcttcatgccca 443

Qy    788 gccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcggcaggcagt 847
          |||||||||  ||||||  |||||||||  | |||||||  ||||||||||  ||||  ||||
Db    444 cccaggcctctggcaccctcagtgcagcccatgattgggaactcaccatcagcagtcagt 503

Qy    848 ggttcggtttaagagatggcattagagggagcccagtctggatgtggacttggatgccct 907
          || ||||||||||||| |||   |||||||  |   |||  ||||||||||||||||||||
Db    504 ggctcggtttaagagagggccgcagagggaactgggtcctgatgtggacttggatgccct 563

Qy    908 gtgggtatcagttctgctgacactttggcccgaaatagatccagtgctgagcaagcaatg 967
          | ||| || || ||||||||||||  |||||  ||||||  |   ||||||||||||||  ||
Db    564 gggggat-agntctgctgacactgtggcctgaaatanaaaaagtgctgagcaagcagtg 622

Qy    968 tacaccggagcctcagtgagcccatctgcacagtggggagcatggagggatgggt 1022
          ||  | ||||||||||| || ||||||||||| ||||||    |   ||||||  |
Db    623 tatgctggagcctcagt-agaccatctgcacaatggggacgtggagaggatggtt 676
```

RESULT 4
AAA89670
ID AAA89670 standard; cDNA; 561 BP.
XX
AC AAA89670;
XX
DT 08-JAN-2001 (first entry)
XX
DE Human T cell leukaemia/lymphoma 1 nucleotide sequence.
XX
KW Exocytosis pathway protein; Exo protein; antiparkinsonian; antidiabetic;
KW antiallergic; antiasthmatic; nootropic; neuroprotective; anticonvulsant;
KW vulnerary; asthma; inflammation; allergy; Chediak-Higashi syndrome; CHS;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; diabetes;
KW digestion disorder; wound healing disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200043419-A2.
XX
PD 27-JUL-2000.
XX
PF 20-JAN-2000; 2000WO-US01431.
XX
PR 20-JAN-1999; 99US-0116534.
PR 26-JAN-1999; 99US-0117274.
PR 26-JAN-1999; 99US-0117308.
PR 26-JAN-1999; 99US-0117309.
PR 26-JAN-1999; 99US-0117312.
PR 01-FEB-1999; 99US-0118177.
PR 01-FEB-1999; 99US-0118178.
PR 01-FEB-1999; 99US-0118179.
PR 09-FEB-1999; 99US-0119286.
PR 11-FEB-1999; 99US-0119998.
PR 11-FEB-1999; 99US-0119759.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y;
XX
DR WPI; 2000-482908/42.
XX
PT New nucleic acids encoding Exo proteins which are useful in the
PT diagnosis, treatment or prevention of exocytosis-mediated disorders
PT such as asthma, inflammation and allergies -
XX
PS Disclosure; Page 251; 305pp; English.
XX
CC The present sequence encodes a polypeptide which is associated with
CC the exocytosis pathway. cDNA molecules encoding proteins involved in
CC exocytosis have been isolated by yeast one-hybrid and two-hybrid
CC screening. Novel proteins, termed Exo proteins, have been identified that
CC interact with known exocytosis-associated proteins such as GS27, alpha
CC snap, unc18-1, vamp3, snap-23, and the rab family of proteins.
CC Exo proteins and their agonists and antagonists are useful in the
CC diagnosis, treatment or prevention of exocytosis-mediated disorders
CC such as asthma, inflammation, allergies, Chediak-Higashi Syndrome
CC (CHS), Alzheimer's disease, Parkinson's disease, Huntington's disease,
CC diabetes, digestion disorders and wound healing disorders.
CC The nucleic acids, antagonists or agonists of Exo proteins are useful
CC in gene therapy. The nucleic acids are also useful for generating
CC transgenic or knock-out animals which can be used in the
CC development and screening of therapeutically useful reagents.
XX
SQ Sequence 561 BP; 110 A; 160 C; 151 G; 135 T; 5 other;

Query Match 5.5%; Score 62.8; DB 21; Length 561;
Best Local Similarity 57.7%; Pred. No. 3.4e-07;
Matches 112; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```
Qy    208 gtgcacttgtggcagatggcagtgcatacccgggagctactctcctccggccagatgccc 267
          |||| |||| | | |   |  ||        ||  ||   |  || ||  ||||||   |
Db     87 gtgctcttgcgtcgggaagacgtcgtcctggggaggcctatgacccccacccagataggc 146

Qy    268 ttctcccagctgcccgccgtgtggcagctctaccccgggaggaagtaccgagcagcggat 327
              || ||||||   | ||||||||||||||||| |   |    ||||||  |  | ||
Db    147 ccaagcctgctgcctatcatgtggcagctctaccctgatggacgataccgatcctcagac 206

Qy    328 tccagtttctgggaaatagcagaccatggccagattgactctatggagcagctggtccta 387
          ||||||||||||    |||   ||||   | ||||||||    |||||  |  || | ||
Db    207 tccagtttctggcgcttagtgtaccacatcaagattgacggcgtggaggacatgcttctc 266

Qy    388 acatatcagccgga 401
              | ||| ||
Db    267 gagctgctgccaga 280
```

RESULT 5
AAI60765
ID AAI60765 standard; cDNA; 1235 BP.
XX
AC AAI60765;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4754.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.

```
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM41609.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 4754; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 1235 BP; 259 A; 359 C; 330 G; 287 T; 0 other;


  Query Match             5.5%;  Score 62.8;  DB 22;  Length 1235;
  Best Local Similarity   57.7%;  Pred. No. 4.7e-07;
  Matches  112;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy    208 gtgcacttgtggcagatggcagtgcatacccgggagctactctcctccggccagatgccc 267
          |||| |||| | | |   |  ||        ||  ||   |  || ||  ||||||   |
Db    192 gtgctcttgcgtcgggaagacgtcgtcctggggaggcctatgacccccacccagataggc 251

Qy    268 ttctcccagctgcccgccgtgtggcagctctaccccgggaggaagtaccgagcagcggat 327
              || ||||||   | ||||||||||||||| |   |    |||||| |  | ||
Db    252 ccaagcctgctgcctatcatgtggcagctctaccctgatggacgataccgatcctcagac 311

Qy    328 tccagtttctgggaaatagcagaccatggccagattgactctatggagcagctggtccta 387
          ||||||||||||    |||   |||| | ||||||||    ||||| |  || | ||
Db    312 tccagtttctggcgcttagtgtaccacatcaagattgacggcgtggaggacatgcttctc 371

Qy    388 acatatcagccgga 401
              | ||| ||
Db    372 gagctgctgccaga 385


RESULT   6
AAT18876
ID   AAT18876 standard; cDNA; 1324 BP.
XX
AC   AAT18876;
XX
DT   23-AUG-1996  (first entry)
XX
DE   Human TCL-1 cDNA.
XX
KW   TCL-1; chromosome-14; leukaemia; lymphoma; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             49..390
FT                   /*tag=  a
XX
PN   WO9613514-A1.
XX
PD   09-MAY-1996.
XX
PF   23-OCT-1995;   95WO-US13663.
XX
PR   27-OCT-1994;   94US-0330272.
XX
PA   (RAGG-) RAGGIO-ITALGENE SPA.
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Croce CM,  Russo G;
XX
DR   WPI; 1996-239444/24.
DR   P-PSDB; AAR94974.
XX
PT   New TCL-1 protein and gene associated with chromosome 14
PT   abnormalities - useful to develop prods. for detection, treatment
PT   and prevention of diseases such as T-cell leukaemia(s) and
PT   lymphoma(s)
XX
PS   Claim 4; Page 66-67; 105pp; English.
XX
CC   A cDNA clone (AAT18876) codes for human TCL-1 protein (AAR94974).
CC   It was obtd. from the ALL1 pre-B-cell line cDNA library by screening
CC   with probes derived from the human genomic library P1 prepd. from
CC   human placental DNA. The cDNA was incorporated into vector pQE30 to
CC   allow expression in Escherichia coli.  A TCL-1 genomic sequence
CC   (AAT18877) has also been identified.  The TCL-1 gene is expressed at
CC   high levels in leukaemic cells carrying a t(14:14)(q11;q32)
CC   translocation or an inv(14)(q11;q32) inversion.  The gene and its
CC   prods. can be used for the diagnosis or treatment of conditions
CC   associated with chromosome 14 abnormalities, esp. T-cell leukaemia
CC   and lymphoma.
XX
SQ   Sequence 1324 BP; 293 A; 376 C; 348 G; 307 T; 0 other;


  Query Match             5.5%;  Score 62.8;  DB 17;  Length 1324;
  Best Local Similarity   57.7%;  Pred. No. 4.9e-07;
  Matches  112;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy    208 gtgcacttgtggcagatggcagtgcatacccgggagctactctcctccggccagatgccc 267
          |||| |||| | | |   |  ||        ||  ||   |  || ||  ||||||   |
Db    190 gtgctcttgcgtcgggaagacgtcgtcctggggaggcctatgacccccacccagataggc 249

Qy    268 ttctcccagctgcccgccgtgtggcagctctaccccgggaggaagtaccgagcagcggat 327
              || ||||||   | ||||||||||||||| |   |    |||||| |  | ||
Db    250 ccaagcctgctgcctatcatgtggcagctctaccctgatggacgataccgatcctcagac 309

Qy    328 tccagtttctgggaaatagcagaccatggccagattgactctatggagcagctggtccta 387
          ||||||||||||    |||   |||| | ||||||||    ||||| |  || | ||
Db    310 tccagtttctggcgcttagtgtaccacatcaagattgacggcgtggaggacatgcttctc 369

Qy    388 acatatcagccgga 401
              | ||| ||
Db    370 gagctgctgccaga 383


RESULT   7
AAI58979
ID   AAI58979 standard; cDNA; 1368 BP.
XX
AC   AAI58979;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 1182.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
```

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM39823.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 1182; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1368 BP; 287 A; 395 C; 366 G; 320 T; 0 other;

Query Match 5.5%; Score 62.8; DB 22; Length 1368;
Best Local Similarity 57.7%; Pred. No. 4.9e-07;
Matches 112; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 208 gtgcacttgtggcagatggcagtgcatacccgggagctactctcctccggccagatgccc 267
Db 247 gtgctcttgcgtcgggaagacgtcgtcctggggaggcctatgacccccacccagataggc 306

Qy 268 ttctcccagctgcccgccgtgtggcagctctacccgggaggaagtaccgagcagcggat 327
Db 307 ccaagcctgctgcctatcatgtggcagctctaccctgatggacgataccgatcctcagac 366

Qy 328 tccagtttctgggaaatagcagaccatggccagattgactctatggagcagctggtccta 387
Db 367 tccagtttctggcgcttagtgtaccacatcaagattgacggcgtggaggacatgcttctc 426

Qy 388 acatatcagccgga 401
Db 427 gagctgctgccaga 440

RESULT 8
AAT18877
ID AAT18877 standard; DNA; 4922 BP.
XX
AC AAT18877;
XX
DT 23-AUG-1996 (first entry)
XX
DE Human TCL-1 gene.
XX
KW TCL-1; chromosome-14; leukaemia; lymphoma; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT TATA_signal 422..426
FT /*tag= a
FT exon 462..627
FT /*tag= b
FT /label= Exon-I
FT intron 628..2203
FT /*tag= c
FT /label= Intron-I
FT exon 2204..2380
FT /*tag= d
FT /label= Exon-II
FT intron 2381..2799
FT /*tag= e
FT /label= Intron-II
FT exon 2800..2853
FT /*tag= f
FT /label= Exon-III
FT intron 2854..3726
FT /*tag= g
FT /label= Intron-III
FT exon 3727..4643
FT /*tag= h
FT /label= Exon-III+3'UTR
XX
PN WO9613514-A1.
XX
PD 09-MAY-1996.
XX
PF 23-OCT-1995; 95WO-US13663.
XX
PR 27-OCT-1994; 94US-0330272.
XX
PA (RAGG-) RAGGIO-ITALGENE SPA.
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Croce CM, Russo G;
XX
DR WPI; 1996-239444/24.
DR P-PSDB; AAR94975.
XX
PT New TCL-1 protein and gene associated with chromosome 14
PT abnormalities - useful to develop prods. for detection, treatment
PT and prevention of diseases such as T-cell leukaemia(s) and
PT lymphoma(s)
XX
PS Disclosure; Page 69-72; 105pp; English.
XX
CC The human TCL-1 gene (AAT18877) was isolated from the region between
CC 2 sets of breakpoints of approx. 160 kb on the TCL-1 locus of
CC chromosome 14. The TCL-1 gene is expressed at high levels in
CC leukaemic cells carrying a t(14:14)(q11;q32) translocation or an
CC inv(14)(q11;q32) inversion. The gene, probes derived from it,
CC antisense sequences, and the encoded protein (AAR94975) can be used
CC for the diagnosis or treatment of conditions associated with

```
CC   increased expression of TCL-1 proteins and/or with chromosome 14
CC   abnormalities, esp. T-cell leukaemia and lymphoma.
XX
SQ   Sequence 4922 BP; 1091 A; 1278 C; 1168 G; 1202 T; 183 other;


  Query Match             4.3%;  Score 50;  DB 17;  Length 4922;
  Best Local Similarity   57.4%;  Pred. No. 0.0023;
  Matches   89;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy     208 gtgcacttgtggcagatggcagtgcatacccgggagctactctcctccggccagatgccc 267
           |||| |||| | | |   |  ||        ||  ||   |  || ||  | ||||   |
Db    2228 gtgctcttgcgtcgggaagacgtcgtcctggggaggcctatgacccccaccnagataggc 2287

Qy     268 ttctcccagctgcccgccgtgtggcagctctaccccgggaggaagtaccgagcagcggat 327
             || ||||||   | ||||||||||||||||| |   |    |||||| |  | ||
Db    2288 ccaagcctgctgcctatcatgtggcagctctaccctgatggacgataccgatcctcagac 2347

Qy     328 tccagtttctgggaaatagcagaccatggccagat 362
           ||||||||||||    |||   ||||   | || |
Db    2348 tccagtttctggcgcttagtgtaccacatcaaggt 2382


RESULT   9
AAA10594/c
ID   AAA10594 standard; DNA; 10732 BP.
XX
AC   AAA10594;
XX
DT   29-JUN-2000  (first entry)
XX
DE   Gene encoding a subunit of cellulose synthase.
XX
KW   Cellulose synthase; cellulose production; increase yield; ds.
XX
OS   Vigna angularis.
XX
PN   JP2000060568-A.
XX
PD   29-FEB-2000.
XX
PF   26-AUG-1998;   98JP-0239998.
XX
PR   26-AUG-1998;   98JP-0239998.
XX
PA   (MIZU/) MIZUNO K.
PA   (OJIP ) OJI PAPER CO.
XX
DR   WPI; 2000-342371/30.
DR   P-PSDB; AAY85179.
XX
PT   A gene encoding a cellulose synthetic equipment - for the improvement
PT   in the amount of cellulose synthesised in a plant body
XX
PS   Claim 2; Page 14-21; 32pp; Japanese.
XX
CC   This sequence represents a gene encoding a subunit of the cellulose
CC   synthase complex of Vigna angularis. The invention relates to subunits of
CC   cellulose synthetic equipment, that can be used to increase the amount of
CC   cellulose synthesised by a plant. The proteins and genes encoding them
CC   can also be used to improve the properties of the cellulose being
CC   produced by a plant.
XX
SQ   Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 2251 other;


  Query Match             4.0%;  Score 45.6;  DB 21;  Length 10732;
  Best Local Similarity   13.2%;  Pred. No. 0.047;
  Matches   68;  Conservative  211;  Mismatches  238;  Indels    0;  Gaps    0;

Qy     419 gggagtggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctc 478
           :::: |: |:   :::: ::  |: ::  :: :| ::    | : :   | |: :  |:
Db   10134 SRSRTTSTCYTAKYSYSTBSTNCYSDSRGSDNSTAKYASTTTYDAYDARCTTDNSTRCYR 10075

Qy     479 agtgaggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcagg 538
           |:| : :: |: : ::    |:::    :: : : :: :     : :|:| : :  |:
Db   10074 ASTYDASRRCYSSTYRASRGTBSDDCCTBSRYSCYSYDASRYANCDTBCYTTBAKYRARC 10015

Qy     539 cagaaggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtcct 598
           ::  |::| :| : : ::   |  :| | :  ::|: :: |:  :||  |: :   :|:
Db   10014 YDAYAKRCNSTSRYRAYSDSTCYTRCRCSRNCYSTYSYSSTYRASTTBTTYCYTCTBCSR 9955

Qy     599 gtcctgtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggcc 658
           : :: ::::    :    :|:  : |   :  :||  :: :  : : :: : : ::| ::
Db    9954 RCYSSRYSSTCNCYSYCCYTSRYSSTTTNSTCRCTTYSYNSTTBTBYSYSSTYSSRGYSS 9895

Qy     659 ctgtctgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcctct 718
             ::: |  :      : |:| : :::  |:  :::: :: ::|::    |:   : | |
Db    9894 RGSDSRGNCYYNSTNCYDASTSDTBYSRCCYTYSYSSTDSTDSTDSTNSTTBSDCYTTTT 9835

Qy     719 gtgcccctcatgctgatcttctagatgccactcccaaatccccttcatacccaccaggat 778
           : :: :    |   ::  :: | :: :   |:  ::  |::   : :   :::   :::
Db    9834 BSRSTSDSTSTYRCSRSDYDATBSDNSTNCCYDASRTBTBSTNCYARCYTBYDARCSRDS 9775

Qy     779 gtgtgcccagccaggcctccagcacccccagtgcagctcgtgattggaaactcaccatcg 838
           :: : : | :  : :: ::  : : : :  : ||::| :|: :::  :||: |: | |
Db    9774 TYSSRGYDANSTSRYSSRYSSTYSSDSTYSAKYCAKSTTBTBCYYDAYDACYDAYDANCY 9715

Qy     839 gcaggcagtggttcggtttaagagatggcattagagggagcccagtctggatgtggactt 898
           :::: : :: :  |  :::| : ::::   ::|  : : : |:: |:| ::::|   ::
Db    9714 SSDSTYTBYCSRRCCCYYDAYSCSRYDARCYDACYSYSNSTCYDATBTTSRYSTTYSYYS 9655

Qy     899 ggatgccctgtgggtatcagttctgctgacactttgg 935
             :   :  |: :| |:    | :||::: :| : :
Db    9654 NCYDATTSRCTBYSTBTBTTSRCAKCTBDSTSTAKNS 9618


RESULT  10
AAV84490
ID   AAV84490 standard; DNA; 1533 BP.
XX
AC   AAV84490;
XX
DT   01-MAR-1999  (first entry)
XX
DE   Human secreted protein gene 80 clone HNFAE54.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
PN   WO9854963-A2.
XX
PD   10-DEC-1998.
XX
PF   04-JUN-1998;   98WO-US11422.
XX
PR   18-DEC-1997;   97US-0070923.
PR   06-JUN-1997;   97US-0048877.
PR   06-JUN-1997;   97US-0048881.
PR   06-JUN-1997;   97US-0048884.
PR   06-JUN-1997;   97US-0048893.
PR   06-JUN-1997;   97US-0048896.
PR   06-JUN-1997;   97US-0048899.
PR   06-JUN-1997;   97US-0048915.
PR   06-JUN-1997;   97US-0048949.
PR   06-JUN-1997;   97US-0048964.
PR   06-JUN-1997;   97US-0048972.
PR   06-JUN-1997;   97US-0049020.
```

PR 06-JUN-1997; 97US-0049375.
PR 05-SEP-1997; 97US-0057628.
PR 05-SEP-1997; 97US-0057635.
PR 05-SEP-1997; 97US-0057644.
PR 05-SEP-1997; 97US-0057647.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057661.
PR 05-SEP-1997; 97US-0057667.
PR 05-SEP-1997; 97US-0057761.
PR 05-SEP-1997; 97US-0057764.
PR 05-SEP-1997; 97US-0057770.
PR 05-SEP-1997; 97US-0057775.
PR 05-SEP-1997; 97US-0057778.
PR 06-JUN-1997; 97US-0048875.
PR 06-JUN-1997; 97US-0048878.
PR 06-JUN-1997; 97US-0048882.
PR 06-JUN-1997; 97US-0048885.
PR 06-JUN-1997; 97US-0048894.
PR 06-JUN-1997; 97US-0048897.
PR 06-JUN-1997; 97US-0048900.
PR 06-JUN-1997; 97US-0048916.
PR 06-JUN-1997; 97US-0048962.
PR 06-JUN-1997; 97US-0048970.
PR 06-JUN-1997; 97US-0048974.
PR 06-JUN-1997; 97US-0049373.
PR 05-SEP-1997; 97US-0057584.
PR 05-SEP-1997; 97US-0057629.
PR 05-SEP-1997; 97US-0057642.
PR 05-SEP-1997; 97US-0057645.
PR 05-SEP-1997; 97US-0057648.
PR 05-SEP-1997; 97US-0057651.
PR 05-SEP-1997; 97US-0057662.
PR 05-SEP-1997; 97US-0057668.
PR 05-SEP-1997; 97US-0057762.
PR 05-SEP-1997; 97US-0057765.
PR 05-SEP-1997; 97US-0057771.
PR 05-SEP-1997; 97US-0057776.
PR 06-JUN-1997; 97US-0048876.
PR 06-JUN-1997; 97US-0048880.
PR 06-JUN-1997; 97US-0048883.
PR 06-JUN-1997; 97US-0048892.
PR 06-JUN-1997; 97US-0048895.
PR 06-JUN-1997; 97US-0048898.
PR 06-JUN-1997; 97US-0048901.
PR 06-JUN-1997; 97US-0048917.
PR 06-JUN-1997; 97US-0048963.
PR 06-JUN-1997; 97US-0048971.
PR 06-JUN-1997; 97US-0049019.
PR 06-JUN-1997; 97US-0049374.
PR 05-SEP-1997; 97US-0057627.
PR 05-SEP-1997; 97US-0057634.
PR 05-SEP-1997; 97US-0057643.
PR 05-SEP-1997; 97US-0057646.
PR 05-SEP-1997; 97US-0057649.
PR 05-SEP-1997; 97US-0057654.
PR 05-SEP-1997; 97US-0057666.
PR 05-SEP-1997; 97US-0057760.
PR 05-SEP-1997; 97US-0057763.
PR 05-SEP-1997; 97US-0057769.
PR 05-SEP-1997; 97US-0057774.
PR 05-SEP-1997; 97US-0057777.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA;
PI Fan P, Feng P, Ferrie AM, Fischer CL, Florence C;
PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DW;
PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM;
PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
XX
DR WPI; 1999-059865/05.
DR P-PSDB; AAW88613.
XX
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Claim 4; Page 343; 772pp; English.
XX
CC The invention relates to nucleic acid sequences (AAV84411 to AAV84633)
CC encoding human secreted proteins (AAW88534 to AAW88756). The secreted
CC protein gene sequences are deposited with the ATCC under deposit numbers
CC ATCC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC cells comprising recombinant vectors containing the nucleic acid
CC sequences are used for the recombinant production of the secreted
CC proteins. The polynucleotide and amino acid sequences are useful for are
CC useful for preventing, treating or ameliorating medical conditions e.g.
CC by protein or gene therapy. Pathological conditions can be also
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, neurodegenerative
CC disorders, developmental abnormalities and foetal deficiencies, blood
CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC such as osteoporosis, arthritis or malignancies, diseases of testes,
CC lung or thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The present sequence represents a gene encoding a human secreted protein
CC (see descriptor line for gene number and clone identification).
XX
SQ Sequence 1533 BP; 272 A; 454 C; 495 G; 302 T; 10 other;

Query Match 3.6%; Score 41.6; DB 20; Length 1533;
Best Local Similarity 53.0%; Pred. No. 0.25;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
Qy    985 gagcccatctgcacagtggggagcatggagggatgggtttggcctgtgcttctgcttatt 1044
          ||| | ||     | ||  ||    ||  || |   | || | |    | |||   |
Db   1354 gaggctatgctttgggggacaggccagaatgaagtcctgggtcaggagcccagctggct 1413

Qy   1045 cagtccttcagctcacggaagggatgctagtccgtgaaggtgacctcacagtactggtta 1104
            | ||  | ||  | | ||||  | |||||  ||   |   || |||  ||    |
Db   1414 gggcccagctgcctatgtaaggccttctagtttgttctgagacccccaccccacgaacca 1473

Qy   1105 attaaactttattgctcactgtcaaaaaaaaaaaaaaaaaaaaaaaaaa 1152
          | |  |  | |     || |  ||||||||||||||||||||||||||
Db   1474 aatccaaataaagtgacattcccaaaaaaaaaaaaaaaaaaaaaaaaaa 1521
```

RESULT 11
AAD04477
ID AAD04477 standard; cDNA; 1719 BP.
XX
AC AAD04477;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human sphingosine kinase type 1 (hsk1) cDNA.
XX
KW Human; sphingosine kinase type 1; sk1; chromosome 17q25.2;
KW sphingosine-1-phosphate; S1P; drug screening; therapy; haemostasis;
KW thrombosis; allergic reaction; proliferative disease; cancer;
KW haematopoietic disorder; leukaemia; cardiovascular disease; stroke;
KW atherosclerosis; coronary artery disease; dyslipidaemia; diabetes;
KW autoimmune disease; inflammatory disease; multiple sclerosis;
KW T helper-1 related disease; chronic obstructive pulmonary disease;
KW asthma; myocardial infarction; neurodegenerative disorder;
KW wound healing; embryogenesis; anticoagulant; cerebroprotective;
KW neuroprotective; antipsoriatic; antiarthritic; cytostatic; cardiant;
KW vulnerary; ss.

```
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             270..1424
FT                   /*tag=  a
FT                   /product= "Human sphingosine kinase type 1 (hsk1)"
FT                   /note= "CDS is specifically claimed in claim 2 and
FT                   shown as SEQ ID NO 2"
FT   polyA_signal    1675..1681
FT                   /*tag=  b
FT   misc_feature    264..273
FT                   /*tag=  c
FT                   /note= "Translational initiator ATG is in a partial
FT                   Kozak consensus"
XX
PN   WO200131029-A2.
XX
PD   03-MAY-2001.
XX
PF   27-OCT-2000; 2000WO-EP09498.
XX
PR   28-OCT-1999;   99US-0162307.
PR   07-FEB-2000; 2000US-0180525.
XX
PA   (WARN ) WARNER LAMBERT CO.
XX
PI   Allen J,  Gosink M,  Melendez AJ,  Takacs L;
XX
DR   WPI; 2001-300510/31.
DR   P-PSDB; AAE00924.
XX
PT   New human sphingosine kinase type I gene for screening drug candidates
PT   particularly inhibitors used for preventing or treating e.g.
PT   atherosclerosis, thrombosis, asthma and diabetes -
XX
PS   Claim 2; Fig 1; 91pp; English.
XX
CC   The present sequence is human sphingosine kinase type 1 (hsk1) cDNA.
CC   The hsk1 gene is located on chromosome 17q25.2. The sk1 converts the
CC   substrate sphingosine to sphingosine-1-phosphate (S1P). The sk1 gene
CC   and encoded polypeptide are applicable in screening drug candidates
CC   particularly inhibitors for preventing or treating disorders such as
CC   haemostasis, thrombosis, allergic reactions, proliferative diseases
CC   including cancer, haematopoietic disorders such as leukaemia,
CC   cardiovascular diseases such as stroke, atherosclerosis and coronary
CC   artery disease, dyslipidaemia, diabetes including type I and type II
CC   diabetes, autoimmune and inflammatory diseases such as multiple
CC   sclerosis, T helper-1 related diseases, chronic obstructive pulmonary
CC   disease, asthma, myocardial infarction, neurodegenerative disorders,
CC   natural wound healing processes and embryogenesis.
XX
SQ   Sequence 1719 BP; 305 A; 529 C; 556 G; 329 T; 0 other;


  Query Match             3.6%;  Score 41.6;  DB 22;  Length 1719;
  Best Local Similarity   53.0%;  Pred. No. 0.26;
  Matches   89;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy     985 gagcccatctgcacagtggggagcatggagggatgggtttggcctgtgcttctgcttatt 1044
           ||| | ||       | ||  ||    ||  || |   | || | |     | |||   |
Db    1550 gaggctatgctttgggggacaggccagaatgaagtcctgggtcaggagcccagctggct 1609

Qy    1045 cagtccttcagctcacggaagggatgctagtccgtgaaggtgacctcacagtactggtta 1104
             | ||  | ||  | | ||||  | |||||  ||   |    || |||   ||     |
Db    1610 gggcccagctgcctatgtaaggccttctagtttgttttgagacccccaccccacgaacca 1669

Qy    1105 attaaactttattgctcactgtcaaaaaaaaaaaaaaaaaaaaaaaaa 1152
           | |  |  | |     || |  ||||||||||||||||||||||||||
Db    1670 aatccaaataaagtgacattcccaaaaaaaaaaaaaaaaaaaaaaaaa 1717


RESULT  12
```

```
AAC75966
ID   AAC75966 standard; cDNA; 1512 BP.
XX
AC   AAC75966;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF1521 polynucleotide sequence SEQ ID NO:3041.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB41757.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 5; Page 2268-2269; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence 1512 BP; 323 A; 408 C; 488 G; 292 T; 1 other;
```

```
  Query Match             3.6%;  Score 41;  DB 21;  Length 1512;
  Best Local Similarity   89.8%;  Pred. No. 0.35;
  Matches   44;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy   1104 aattaaactttattgctcactgtcaaaaaaaaaaaaaaaaaaaaaaaa 1152
          ||| | | |||| |||||||||| ||||||||||||||||||||||||
Db   1454 aataataatttaatgctcactgtgaaaaaaaaaaaaaaaaaaaaaaaa 1502


RESULT  13
AAC77441
ID   AAC77441 standard; cDNA; 2440 BP.
XX
AC   AAC77441;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF2996 polynucleotide sequence SEQ ID NO:5991.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB43232.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 5; Page 5164-5166; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence 2440 BP; 551 A; 706 C; 787 G; 394 T; 2 other;


  Query Match             3.5%;  Score 40.4;  DB 21;  Length 2440;
  Best Local Similarity   88.0%;  Pred. No. 0.63;
  Matches   44;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy   1103 taattaaactttattgctcactgtcaaaaaaaaaaaaaaaaaaaaaaaa 1152
          |||| || |||||||| |||||  ||||||||||||||||||||||||
Db   2378 taataaatctttattgagcactgaaaaaaaaaaaaaaaaaaaaaaaa 2427


RESULT  14
AAH34812
ID   AAH34812 standard; cDNA; 2617 BP.
XX
AC   AAH34812;
XX
DT   03-SEP-2001  (first entry)
XX
DE   Human colon cancer antigen encoding cDNA SEQ ID NO:1894.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   colorectal carcinoma; chromosome 7; ss.
XX
OS   Homo sapiens.
XX
PN   WO200122920-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26524.
XX
PR   29-SEP-1999;   99US-0157137.
PR   03-NOV-1999;   99US-0163280.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Barash SC,  Birse CE,  Rosen CA;
XX
DR   WPI; 2001-235357/24.
DR   P-PSDB; AAG75407.
XX
PT   Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT   useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS   Claim 1; Page 3407-3408; 9803pp; English.
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer-associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate P
CC   expression. For example, N and P may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of P by expressing
CC   inactive proteins or to supplement the patients own production of P.
CC   Additionally, N may be used to produce the colon cancer-associated Ps,
CC   by inserting the nucleic acids into a host cell and culturing the cell
CC   to express the proteins. N and P can be used in the prevention, diagnosis
CC   and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
```

```
CC   and AAB77789 represent sequences used in the exemplification of the
CC   present invention.
CC   N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC   missing at time of publication, meaning no sequences are present for
CC   SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ   Sequence 2617 BP; 663 A; 672 C; 685 G; 595 T; 2 other;


  Query Match             3.5%;  Score 39.8;  DB 22;  Length 2617;
  Best Local Similarity   50.2%;  Pred. No. 0.93;
  Matches  121;  Conservative    1;  Mismatches  118;  Indels    1;  Gaps    1;

Qy    421 gagtggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccctcag 480
          :|| |||||| |     ||  | ||| ||||  ||||||  ||  |   | | | |   |
Db     88 kagcggctgggcgctgaggatcagccgcttcctgcctggattccacagcttcgcgccgtg 147

Qy    481 tgaggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgcaggca 540
          | |   ||  |   | ||||| |  |        ||  |||| |  ||   |||||||| 
Db    148 t-actgtcgcccatccctgcgcgcccagcctgccaagcagcgtgccccggttgcaggcg 206

Qy    541 gaaggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtcctgt 600
            | ||||  ||  |    ||| ||| |   | ||    |  ||  | |     | ||
Db    207 tcatgcagcgggcgcgacccacgctctgggccgctgcgctgactctgctggtgctgctcc 266

Qy    601 cctgtcccactcaggtctgtacttagggcagctggcctggatgggcttcactggggccct 660
          | | ||  |   ||   |  ||   |  |||| | | || | || | |  ||| || ||
Db    267 gcgggccgccggtggcgcgggctggcgcgagctcggcgggcttgggtcccgtggtgcgct 326

Qy    661 g 661
          |
Db    327 g 327


RESULT  15
AAA77669
ID   AAA77669 standard; cDNA; 693 BP.
XX
AC   AAA77669;
XX
DT   07-NOV-2000  (first entry)
XX
DE   Human PRO1286 cDNA sequence SEQ ID NO:198.
XX
KW   Human; PRO; promotion; inhibition; angiogenesis; cardiovascularisation;
KW   diagnosis; trauma; wound; cancer; atherosclerosis; cardiac hypertrophy;
KW   angiogenic; proliferative; cardiant; cardiovascular; antiatherosclerotic;
KW   cytostatic; gene therapy; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200032221-A2.
XX
PD   08-JUN-2000.
XX
PF   30-NOV-1999;   99WO-US28313.
XX
PR   01-DEC-1998;   98WO-US25108.
PR   16-DEC-1998;   98US-0112850.
PR   12-JAN-1999;   99US-0115554.
PR   08-MAR-1999;   99WO-US05028.
PR   12-MAR-1999;   99US-0123957.
PR   28-APR-1999;   99US-0131445.
PR   14-MAY-1999;   99US-0134287.
PR   02-JUN-1999;   99WO-US12252.
PR   23-JUN-1999;   99US-0141037.
PR   20-JUL-1999;   99US-0144758.
PR   26-JUL-1999;   99US-0145698.
PR   01-SEP-1999;   99WO-US20111.
PR   08-SEP-1999;   99WO-US20594.
PR   13-SEP-1999;   99WO-US20944.
PR   15-SEP-1999;   99WO-US21090.
PR   15-SEP-1999;   99WO-US21547.
PR   05-OCT-1999;   99WO-US23089.
PR   29-OCT-1999;   99US-0162506.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Baker KP,  Ferrara N,  Gerber H,  Hillan KJ,  Goddard A;
PI   Godowski PJ,  Gurney AL,  Klein RD,  Kuo SS,  Paoni NF,  Smith V;
PI   Watanabe CK,  Williams PM,  Wood WI;
XX
DR   WPI; 2000-412154/35.
DR   P-PSDB; AAB24426.
XX
PT   Nucleic acids encoding PRO polypeptides useful for preventing,
PT   diagnosing and treating diagnosing a cardiovascular, endothelial or
PT   angiogenic disorders in mammals -
XX
PS   Claim 61; Fig 77; 315pp; English.
XX
CC   The present invention describes nucleic acids encoding PRO polypeptides
CC   useful for preventing, diagnosing and treating diagnosing a
CC   cardiovascular, endothelial or angiogenic disorder in mammals by
CC   modulating cell proliferation, angiogenesis and cardiovascularisation,
CC   and for identifying agonists and antagonists of these processes. The
CC   nucleic acids and the proteins they encode may be used in the
CC   prevention, treatment and diagnosis of diseases associated with
CC   inappropriate PRO expression such as cardiovascular, endothelial or
CC   angiogenic disorders in mammals (e.g. atherosclerosis, cancers and
CC   cardiac hypertrophy). For example, the nucleic acids (NCs) and vectors
CC   containing them and the PRO polypeptide may be used to treat disorders
CC   associated with decreased PRO expression. AAA77510 to AAA77721 and
CC   AAB24388 to AAB24435 represent nucleotide and protein sequences used in
CC   the exemplification of the present invention.
XX
SQ   Sequence 693 BP; 184 A; 179 C; 230 G; 100 T; 0 other;


  Query Match             3.4%;  Score 39.2;  DB 21;  Length 693;
  Best Local Similarity   62.0%;  Pred. No. 0.78;
  Matches   62;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy   1053 cagctcacggaagggatgctagtccgtgaaggtgacctcacagtactggttaattaaaact 1112
          | |  || ||  | || || |||| |    ||||| || | || ||   | || | |  |
Db    567 cgggccagggccgagacgcgagtcggatgtggtgaactgaaagaaccaataaaatcatgt 626

Qy   1113 ttattgctcactgtcaaaaaaaaaaaaaaaaaaaaaaaaa 1152
          |  |     |     ||||||||||||||||||||||||
Db    627 tcctccaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 666


Search completed: November 29, 2001, 08:59:23
Job time: 17559 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 04:06:24 ; Search time 194.7 Seconds
(without alignments)
1340.024 Million cell updates/sec

Title: US-09-526-329-38
Perfect score: 1152
Sequence: 1 gaggcgggtcccggttgcag.....aaaaaaaaaaaaaaaaaaaa 1152

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	64.4	5.6	7218	1	US-08-232-463-14	Sequence 14, Appl
2	62.8	5.5	1324	2	US-08-330-272-1	Sequence 1, Appli
3	62.8	5.5	1324	5	PCT-US95-13663-1	Sequence 1, Appli
4	50	4.3	4922	2	US-08-330-272-5	Sequence 5, Appli
5	50	4.3	4922	5	PCT-US95-13663-5	Sequence 5, Appli
6	38.8	3.4	1817	1	US-08-473-981A-5	Sequence 5, Appli
7	38.8	3.4	1817	2	US-08-474-087-5	Sequence 5, Appli
8	38.6	3.4	284	4	US-09-040-984-80	Sequence 80, Appl
9	37.6	3.3	2017	4	US-09-436-983-1	Sequence 1, Appli
10	37.2	3.2	1781	1	US-08-314-615-2	Sequence 2, Appli
11	37.2	3.2	1781	1	US-08-314-362-2	Sequence 2, Appli
12	37.2	3.2	1781	1	US-08-433-010-2	Sequence 2, Appli
13	37.2	3.2	1781	1	US-08-245-295-4	Sequence 4, Appli
14	37.2	3.2	1781	1	US-08-481-130-4	Sequence 4, Appli
15	37.2	3.2	1781	1	US-08-656-984A-4	Sequence 4, Appli
16	37.2	3.2	1781	1	US-08-482-882-2	Sequence 2, Appli
17	37.2	3.2	1781	1	US-08-485-604-4	Sequence 4, Appli
18	37.2	3.2	1781	1	US-08-483-389-2	Sequence 2, Appli
19	37.2	3.2	1781	2	US-08-487-113D-2	Sequence 2, Appli
20	37.2	3.2	1781	2	US-08-487-595-4	Sequence 4, Appli
21	37.2	3.2	1781	2	US-08-473-503-2	Sequence 2, Appli
22	37.2	3.2	1781	2	US-08-483-932-2	Sequence 2, Appli
23	37.2	3.2	1781	2	US-08-720-420A-2	Sequence 2, Appli
24	37.2	3.2	1781	3	US-08-714-017-2	Sequence 2, Appli
25	37.2	3.2	1781	3	US-08-863-790-2	Sequence 2, Appli
26	37.2	3.2	1781	3	US-08-475-680-2	Sequence 2, Appli
27	37.2	3.2	1781	3	US-08-296-749-2	Sequence 2, Appli
c 28	37	3.2	289	4	US-09-007-005-17	Sequence 17, Appl
c 29	37	3.2	289	4	US-09-244-796-17	Sequence 17, Appl
c 30	37	3.2	322	4	US-09-385-982-216	Sequence 216, App
c 31	37	3.2	322	4	US-09-385-982-362	Sequence 362, App
32	36.6	3.2	4843	3	US-08-986-485-1	Sequence 1, Appli
33	36.4	3.2	570	1	US-08-469-667-13	Sequence 13, Appl
34	36.4	3.2	570	5	PCT-US95-07289-13	Sequence 13, Appl
35	36.4	3.2	1357	6	5340934-7	Patent No. 5340934
36	36.2	3.1	1394	1	US-07-730-953-1	Sequence 1, Appli
37	36	3.1	1147	1	US-08-665-716-1	Sequence 1, Appli
38	36	3.1	3385	1	US-08-405-392-1	Sequence 1, Appli
39	36	3.1	3385	3	US-08-487-691-1	Sequence 1, Appli
40	36	3.1	3385	3	US-08-666-221B-3	Sequence 3, Appli
41	36	3.1	3385	3	US-08-666-221B-9	Sequence 9, Appli
42	35.6	3.1	902	2	US-08-378-939-11	Sequence 11, Appl
c 43	35.4	3.1	703	4	US-09-313-300-6	Sequence 6, Appli
44	35.4	3.1	989	2	US-08-874-460-1	Sequence 1, Appli
45	35.4	3.1	2684	2	US-08-984-171-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
;   APPLICANT: DORNER, F.
;   APPLICANT: SCHEIFLINGER, F.
;   APPLICANT: FALKNER, F. G.
;   TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
;   NUMBER OF SEQUENCES: 52
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Foley & Lardner
;     STREET: 1800 Diagonal Road, Suite 500
;     CITY: Alexandria
;     STATE: VA
;     COUNTRY: USA
;     ZIP: 22313-0299
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/232,463
;     FILING DATE:
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/07/935,313
;     FILING DATE:
;     APPLICATION NUMBER: EP 91 114 300.6
;     FILING DATE: 26-AUG-1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME: BENT, Stephen A.
;     REGISTRATION NUMBER: 29,768
;     REFERENCE/DOCKET NUMBER: 30472/114 IMMU
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (703)836-9300
;     TELEFAX: (703)683-4109
;     TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 7218 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;     CLONE: pTZgpt-Fls
US-08-232-463-14
```

```
  Query Match             5.6%;  Score 64.4;  DB 1;  Length 7218;
  Best Local Similarity   3.5%;  Pred. No. 1.1e-08;
  Matches   14;  Conservative  233;  Mismatches  149;  Indels    0;  Gaps    0;

Qy    417 ctgggagtggctggccctgctggccctgcctcttctggcctggtgtctcctcatgccccc 476
          |||   || |  ||  ||    | ::: :::::::::  :::  : :::::::: : ::::: 
Db   1043 CTGCAGGTCGAGGGAGCTTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1102

Qy    477 tcagtgaggatcttcatgtacctgctcttctgtttgcacacccagcatagcctccttgca 536
          ::  :     ::::: : : ::: :::::::: ::: : : :::  : :  :::::::: : 
Db   1103 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1162

Qy    537 ggcagaaggcagtagggcccctgcacactcagtttctctcgttttccttagttatcagtc 596
            :       :   :    ::::: : : :::  :::::::: ::::::::  :: ::  :: 
Db   1163 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1222

Qy    597 ctgtcctgtcccactcaggtctgtacttagggcagctggcctggatgggcttcactgggg 656
          :: :::: :::: :::   ::: : :::    :  ::  :::   :   :::: :: 
Db   1223 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1282

Qy    657 ccctgtctgtgtgctgagccagtttcccctgctggctgcaagctgtgggttctttctcct 716
          :::: ::: : : ::   ::  :::::::: ::  :: :   :: :   :::::::::: 
Db   1283 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1342

Qy    717 ctgtgcccctcatgctgatcttctagatgccactcccaaatccccttcatacccaccagg 776
          :: : :::::: : ::  ::::::   : :: :::::   :::::::: : ::: :: 
Db   1343 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1402

Qy    777 atgtgtgcccagccaggcctccagcacccccagtgc 812
            : : : :::  ::   ::::  : ::::: || |
Db   1403 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTAC 1438


RESULT   2
US-08-330-272-1
; Sequence 1, Application US/08330272
; Patent No. 5985598
;  GENERAL INFORMATION:
;    APPLICANT:  Russo et al
;    TITLE OF INVENTION:  TCL-1 Gene and Protein and Related
;    TITLE OF INVENTION:  Methods and Compositions
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/330,272
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  6754-027
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 790-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1324 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  49..387
US-08-330-272-1


  Query Match             5.5%;  Score 62.8;  DB 2;  Length 1324;
  Best Local Similarity   57.7%;  Pred. No. 1.3e-08;
  Matches  112;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy    208 gtgcacttgtggcagatggcagtgcatacccgggagctactctcctccggccagatgccc 267
          |||| |||| | | |   |  ||         ||  ||   |  || ||  ||||||   | 
Db    190 GTGCTCTTGCGTCGGGAAGACGTCGTCCTGGGGAGGCCTATGACCCCCACCCAGATAGGC 249

Qy    268 ttctcccagctgcccgccgtgtggcagctctaccccgggaggaagtaccgagcagcggat 327
              || ||||||   | ||||||||||||||| |   |    |||||| |  | ||
Db    250 CCAAGCCTGCTGCCTATCATGTGGCAGCTCTACCCTGATGGACGATACCGATCCTCAGAC 309

Qy    328 tccagtttctgggaaatagcagaccatggccagattgactctatggagcagctggtccta 387
          ||||||||||||    |||   ||||   | ||||||||    ||||| |  || | ||
Db    310 TCCAGTTTCTGGCGCTTAGTGTACCACATCAAGATTGACGGCGTGGAGGACATGCTTCTC 369

Qy    388 acatatcagccgga 401
                | ||| ||
Db    370 GAGCTGCTGCCAGA 383


RESULT   3
PCT-US95-13663-1
; Sequence 1, Application PC/TUS9513663
;  GENERAL INFORMATION:
;    APPLICANT:  Russo et al
;    TITLE OF INVENTION:  TCL-1 Gene and Protein and Related
;    TITLE OF INVENTION:  Methods and Compositions
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/13663
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  6754-027
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 790-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1324 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  49..387
PCT-US95-13663-1


  Query Match             5.5%;  Score 62.8;  DB 5;  Length 1324;
```

```
  Best Local Similarity   57.7%;  Pred. No. 1.3e-08;
  Matches  112;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy    208 gtgcacttgtggcagatggcagtgcatacccgggagctactctcctccggccagatgccc 267
          |||| |||| | | |    |  ||         ||  ||   |  || ||  ||||||   |
Db    190 GTGCTCTTGCGTCGGGAAGACGTCGTCCTGGGGAGGCCTATGACCCCCACCCAGATAGGC 249

Qy    268 ttctcccagctgcccgccgtgtggcagctctaccccgggaggaagtaccgagcagcggat 327
              || ||||||   | ||||||||||||||| |   |    |||||| |  | ||
Db    250 CCAAGCCTGCTGCCTATCATGTGGCAGCTCTACCCTGATGGACGATACCGATCCTCAGAC 309

Qy    328 tccagtttctgggaaatagcagaccatggccagattgactctatggagcagctggtccta 387
          ||||||||||||    |||   ||||   | ||||||||    ||||| |  || | ||
Db    310 TCCAGTTTCTGGCGCTTAGTGTACCACATCAAGATTGACGGCGTGGAGGACATGCTTCTC 369

Qy    388 acatatcagccgga 401
                | ||| ||
Db    370 GAGCTGCTGCCAGA 383


RESULT   4
US-08-330-272-5
; Sequence 5, Application US/08330272
; Patent No. 5985598
;  GENERAL INFORMATION:
;    APPLICANT:  Russo et al
;    TITLE OF INVENTION:  TCL-1 Gene and Protein and Related
;    TITLE OF INVENTION:  Methods and Compositions
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/330,272
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  6754-027
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 790-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4922 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA
US-08-330-272-5


  Query Match             4.3%;  Score 50;  DB 2;  Length 4922;
  Best Local Similarity   57.4%;  Pred. No. 0.00013;
  Matches   89;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy    208 gtgcacttgtggcagatggcagtgcatacccgggagctactctcctccggccagatgccc 267
          |||| |||| | | |    |  ||         ||  ||   |  || ||  | ||||   |
Db   2228 GTGCTCTTGCGTCGGGAAGACGTCGTCCTGGGGAGGCCTATGACCCCCACCNAGATAGGC 2287

Qy    268 ttctcccagctgcccgccgtgtggcagctctaccccgggaggaagtaccgagcagcggat 327
              || ||||||   | ||||||||||||||| |   |    |||||| |  | ||

Db   2288 CCAAGCCTGCTGCCTATCATGTGGCAGCTCTACCCTGATGGACGATACCGATCCTCAGAC 2347

Qy    328 tccagtttctgggaaatagcagaccatggccagat 362
          ||||||||||||    |||   ||||   | || |
Db   2348 TCCAGTTTCTGGCGCTTAGTGTACCACATCAAGGT 2382


RESULT   5
PCT-US95-13663-5
; Sequence 5, Application PC/TUS9513663
;  GENERAL INFORMATION:
;    APPLICANT:  Russo et al
;    TITLE OF INVENTION:  TCL-1 Gene and Protein and Related
;    TITLE OF INVENTION:  Methods and Compositions
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/13663
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  6754-027
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 790-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4922 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA
PCT-US95-13663-5


  Query Match             4.3%;  Score 50;  DB 5;  Length 4922;
  Best Local Similarity   57.4%;  Pred. No. 0.00013;
  Matches   89;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy    208 gtgcacttgtggcagatggcagtgcatacccgggagctactctcctccggccagatgccc 267
          |||| |||| | | |    |  ||         ||  ||   |  || ||  | ||||   |
Db   2228 GTGCTCTTGCGTCGGGAAGACGTCGTCCTGGGGAGGCCTATGACCCCCACCNAGATAGGC 2287

Qy    268 ttctcccagctgcccgccgtgtggcagctctaccccgggaggaagtaccgagcagcggat 327
              || ||||||   | ||||||||||||||| |   |    |||||| |  | ||
Db   2288 CCAAGCCTGCTGCCTATCATGTGGCAGCTCTACCCTGATGGACGATACCGATCCTCAGAC 2347

Qy    328 tccagtttctgggaaatagcagaccatggccagat 362
          ||||||||||||    |||   ||||   | || |
Db   2348 TCCAGTTTCTGGCGCTTAGTGTACCACATCAAGGT 2382


RESULT   6
US-08-473-981A-5
; Sequence 5, Application US/08473981A
; Patent No. 5629162
;  GENERAL INFORMATION:
;    APPLICANT:  deFougerolles, Antonin R
;    APPLICANT:  Springer, Timothy A
```

```
;    TITLE OF INVENTION:  METHODS OF IDENTIFYING AGENTS WHICH MODULATE
;    TITLE OF INVENTION:  ICAM-3 BINDING TO LFA-1 (AS AMENDED)
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
;      STREET:  1100 NEW YORK AVENUE, N.W. SUITE 600
;      CITY:  WASHINGTON
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/473,981A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MILLONIG, ROBERT C
;      REGISTRATION NUMBER:  34,395
;      REFERENCE/DOCKET NUMBER:  1011.0560004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1817 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  9..1649
US-08-473-981A-5


  Query Match             3.4%;  Score 38.8;  DB 1;  Length 1817;
  Best Local Similarity   64.4%;  Pred. No. 0.14;
  Matches   58;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy   1063 aagggatgctagtccgtgaaggtgacctcacagtactggttaattaaactttattgctca 1122
          || |   ||  |  | ||   ||| ||||| | | |     ||| || | ||    |||
Db   1670 AAAGTTGGCGGGGGCTTGGCTGTGCCCTCAGATTCCGCACCAATAAAGCCTTCAAACTCC 1729

Qy   1123 ctgtcaaaaaaaaaaaaaaaaaaaaaaaaa 1152
          ||   |||||||||||||||||||||||||
Db   1730 CTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1759


RESULT   7
US-08-474-087-5
; Sequence 5, Application US/08474087
; Patent No. 5891841
;  GENERAL INFORMATION:
;    APPLICANT:  de Fougerolles, Antonin R
;    APPLICANT:  Springer, Timothy A
;    TITLE OF INVENTION:  METHODS OF USING INTERCELLULAR ADHESION MOLECULE-      3 (IC
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
;      STREET:  1100 NEW YORK AVENUE, N.W. SUITE 600
;      CITY:  WASHINGTON
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,087
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/038,990
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/712,879
;      FILING DATE:  11-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MILLONIG, ROBERT C
;      REGISTRATION NUMBER:  34,395
;      REFERENCE/DOCKET NUMBER:  1011.0560003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1817 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  9..1649
US-08-474-087-5


  Query Match             3.4%;  Score 38.8;  DB 2;  Length 1817;
  Best Local Similarity   64.4%;  Pred. No. 0.14;
  Matches   58;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy   1063 aagggatgctagtccgtgaaggtgacctcacagtactggttaattaaactttattgctca 1122
          || |   ||  |  | ||   ||| ||||| | | |     ||| || | ||    |||
Db   1670 AAAGTTGGCGGGGGCTTGGCTGTGCCCTCAGATTCCGCACCAATAAAGCCTTCAAACTCC 1729

Qy   1123 ctgtcaaaaaaaaaaaaaaaaaaaaaaaaa 1152
          ||   |||||||||||||||||||||||||
Db   1730 CTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1759


RESULT   8
US-09-040-984-80
; Sequence 80, Application US/09040984
; Patent No. 6210883
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Wang, TongTong
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS
;    TITLE OF INVENTION:  OF LUNG CANCER
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/040,984
;      FILING DATE:  18-MAR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.456
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-622-4900
;      TELEFAX:  206-282-6031
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  80:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  284 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-040-984-80


  Query Match             3.4%;  Score 38.6;  DB 4;  Length 284;
  Best Local Similarity   81.5%;  Pred. No. 0.058;
  Matches   44;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1099 tggttaattaaactttattgctcactgtcaaaaaaaaaaaaaaaaaaaaaaaaa 1152
          | |  || |||| | |||| || ||||| |||||||||||||||||||||||| |
Db    231 TTGGGAAATAAAATGTATTTCTTACTGTGAAAAAAAAAAAAAAAAAAAAAAAANA 284


RESULT   9
US-09-436-983-1
; Sequence 1, Application US/09436983
; Patent No. 6294343
; GENERAL INFORMATION:
;  APPLICANT:  Mack, David
;  APPLICANT:  Gish, Kurt C
;  APPLICANT:  Wilson, Keith E
;  TITLE OF INVENTION:  NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER,
;  TITLE OF INVENTION:  COMPOSITIONS, AND METHODS OF SCREENING FOR COLORECTAL
;  TITLE OF INVENTION:  CANCER MODULATORS
;  FILE REFERENCE:  A-68431/RMS/DAV
;  CURRENT APPLICATION NUMBER: US/09/436,983
;  CURRENT FILING DATE:  1999-11-09
;  NUMBER OF SEQ ID NOS:  8
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 2017
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-436-983-1


  Query Match             3.3%;  Score 37.6;  DB 4;  Length 2017;
  Best Local Similarity   54.3%;  Pred. No. 0.33;
  Matches   76;  Conservative    0;  Mismatches   64;  Indels    0;  Gaps    0;

Qy   1013 agggatgggtttggcctgtgcttctgcttattcagtccttcagctcacggaagggatgct 1072
          | |||   |  | | || |||||    | ||       ||    ||||   || |  |||
Db   1872 atggaaaagactaggctttgcttagtatcatgtccatgtttccttcacctcagtggagct 1931

Qy   1073 agtccgtgaaggtgacctcacagtactggttaattaaactttattgctcactgtcaaaaa 1132
            |  ||         ||||   |  |  | ||| ||| ||| |  |    |   |||||
Db   1932 tctgagttttatactgctcaagatcgtcataaataaaatttttttctcattgtcaaaaaaa 1991

Qy   1133 aaaaaaaaaaaaaaaaaaa 1152
          ||||||||||||||||||
Db   1992 aaaaaaaaaaaaaaaaaaa 2011


RESULT  10
US-08-314-615-2
; Sequence 2, Application US/08314615
; Patent No. 5525487
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Vazeux, Rosemay
;    TITLE OF INVENTION:  I-CAM Related Protein
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/314,615
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/827,689
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27866/30704
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312)346-5750
;      TELEFAX:  (312)984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1781 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-314-615-2


  Query Match             3.2%;  Score 37.2;  DB 1;  Length 1781;
  Best Local Similarity   63.3%;  Pred. No. 0.4;
  Matches   57;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy   1063 aagggatgctagtccgtgaaggtgacctcacagtactggttaattaaactttattgctca 1122
          || |   ||  |  | ||   ||| ||||| | | |     ||| || | ||    |||
Db   1677 AAAGTTGGCGGGGGCTTGGCTGTGCCCTCAGATTCCGCACCAATAAAGCCTTCAAACTCC 1736

Qy   1123 ctgtcaaaaaaaaaaaaaaaaaaaaaaaa 1152
          |    ||||||||||||||||||||||||||
Db   1737 CAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1766


RESULT  11
US-08-314-362-2
; Sequence 2, Application US/08314362
; Patent No. 5532127
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Vazeux, Rosemay
;    TITLE OF INVENTION:  I-CAM Related Protein
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/314,362
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/894,061
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/827,689
;      FILING DATE:  27-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US
;      FILING DATE:  26-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5532127and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/30918
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312)346-5750
;      TELEFAX:  (312)984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1781 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-314-362-2


  Query Match             3.2%;  Score 37.2;  DB 1;  Length 1781;
  Best Local Similarity   63.3%;  Pred. No. 0.4;
  Matches   57;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy   1063 aagggatgctagtccgtgaaggtgacctcacagtactggttaattaaactttattgctca 1122
          || |   ||  |  | ||   ||| ||||| | | |     ||| || | ||    |||
Db   1677 AAAGTTGGCGGGGGCTTGGCTGTGCCCTCAGATTCCGCACCAATAAAGCCTTCAAACTCC 1736

Qy   1123 ctgtcaaaaaaaaaaaaaaaaaaaaaaaaa 1152
          |    ||||||||||||||||||||||||||
Db   1737 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1766


RESULT  12
US-08-433-010-2
; Sequence 2, Application US/08433010
; Patent No. 5663293
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Vazeux, Rosemay
;    TITLE OF INVENTION:  ICAM-Related Protein
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/433,010
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/009,266
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/827,689
;      FILING DATE:  27-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/889,724
;      FILING DATE:  26-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/894,061
;      FILING DATE:  05-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5663293and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  31218
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312)346-5750
;      TELEFAX:  (312)984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1781 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-433-010-2


  Query Match             3.2%;  Score 37.2;  DB 1;  Length 1781;
  Best Local Similarity   63.3%;  Pred. No. 0.4;
  Matches   57;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy   1063 aagggatgctagtccgtgaaggtgacctcacagtactggttaattaaactttattgctca 1122
          || |   ||  |  | ||   ||| ||||| | | |     ||| || | ||    |||
Db   1677 AAAGTTGGCGGGGGCTTGGCTGTGCCCTCAGATTCCGCACCAATAAAGCCTTCAAACTCC 1736

Qy   1123 ctgtcaaaaaaaaaaaaaaaaaaaaaaaaa 1152
          |    ||||||||||||||||||||||||||
Db   1737 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1766


RESULT  13
US-08-245-295-4
; Sequence 4, Application US/08245295
; Patent No. 5700658
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Kilgannon, Patrick D.
;    TITLE OF INVENTION:  ICAM-4 Materials and Methods
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, Suite 6300
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/245,295
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/827,689
;      FILING DATE:  27-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/889,724
;      FILING DATE:  26-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/894,061
```

```
;      FILING DATE:  05-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/009,266
;      FILING DATE:  22-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/102,852
;      FILING DATE:  05-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27866/32055
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1781 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  16..1659
US-08-245-295-4


  Query Match             3.2%;  Score 37.2;  DB 1;  Length 1781;
  Best Local Similarity   63.3%;  Pred. No. 0.4;
  Matches   57;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy   1063 aagggatgctagtccgtgaaggtgacctcacagtactggttaattaaactttattgctca 1122
          || |   ||  |  | ||   ||| ||||| | | |     ||| || | ||    |||
Db   1677 AAAGTTGGCGGGGGCTTGGCTGTGCCCTCAGATTCCGCACCAATAAAGCCTTCAAACTCC 1736

Qy   1123 ctgtcaaaaaaaaaaaaaaaaaaaaaaaaa 1152
          |    |||||||||||||||||||||||||
Db   1737 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1766


RESULT  14
US-08-481-130-4
; Sequence 4, Application US/08481130
; Patent No. 5702917
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Kilgannon, Patrick D.
;    TITLE OF INVENTION:  ICAM-4 Materials and Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,130
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/827,689
;      FILING DATE:  27-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/889,724
;      FILING DATE:  26-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/894,061
;      FILING DATE:  05-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/009,266
;      FILING DATE:  22-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/102,852
;      FILING DATE:  05-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/245,295
;      FILING DATE:  18-MAY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WILLIAMS, JR. JOSEPH A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32713
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1781 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  16..1659
US-08-481-130-4


  Query Match             3.2%;  Score 37.2;  DB 1;  Length 1781;
  Best Local Similarity   63.3%;  Pred. No. 0.4;
  Matches   57;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy   1063 aagggatgctagtccgtgaaggtgacctcacagtactggttaattaaactttattgctca 1122
          || |   ||  |  | ||   ||| ||||| | | |     ||| || | ||    |||
Db   1677 AAAGTTGGCGGGGGCTTGGCTGTGCCCTCAGATTCCGCACCAATAAAGCCTTCAAACTCC 1736

Qy   1123 ctgtcaaaaaaaaaaaaaaaaaaaaaaaaa 1152
          |    |||||||||||||||||||||||||
Db   1737 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1766


RESULT  15
US-08-656-984A-4
; Sequence 4, Application US/08656984A
; Patent No. 5753502
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Kilgannon, Patrick D.
;    TITLE OF INVENTION:  ICAM-4 Materials and Methods
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/656,984A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/827,689
;      FILING DATE:  27-JAN-1992
```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,295
; FILING DATE: 18-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,604
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, JR. JOSEPH A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/33321
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1781 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..1659
US-08-656-984A-4

Query Match 3.2%; Score 37.2; DB 1; Length 1781;
Best Local Similarity 63.3%; Pred. No. 0.4;
Matches 57; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy   1063 aagggatgctagtccgtgaaggtgacctcacagtactggttaattaaactttattgctca 1122
          || |   ||  |  | ||   ||| ||||| | | |     ||| || | ||    |||
Db   1677 AAAGTTGGCGGGGGCTTGGCTGTGCCCTCAGATTCCGCACCAATAAAGCCTTCAAACTCC 1736

Qy   1123 ctgtcaaaaaaaaaaaaaaaaaaaaaaaa 1152
          |    ||||||||||||||||||||||||||
Db   1737 CAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1766
```

Search completed: November 29, 2001, 08:49:30
Job time: 16986 sec